TESIS DOCTORAL

AÑO 2022

IDENTITY FUSION AND MORALITY: THE ROLE OF ADMIRATION, MORAL BELIEFS ABOUT VIOLENCE, AND VISCERAL RESPONSIBILITY

JUANA FRANCISCA CHINCHILLA CALERO

PROGRAMA DE DOCTORADO EN PSICOLOGÍA

ÁNGEL GÓMEZ JIMÉNEZ
ALEXANDRA VÁZQUEZ BOTANA

The Keeper

Ramón Maiden

Identity fusion and morality: The role of admiration, moral beliefs about violence, and visceral responsibility

PhD Thesis

Autor: Juana Chinchilla
Advisors: Ángel Gómez
Alexandra Vázquez

2022

AKNOWLEDGMENTS

At this stage of my academic and personal development, I cannot help turning my eyes back and see that the path lying behind is long. I have mostly walked through it as the explorers of exotic lands that I admired as a child, full of determination and optimism; and, like them, I have encountered some treasures that have made me incommensurably rich. The most valuable of all of them is no other than the people who have walked with me. I want to dedicate them a few words to recognize that, without their help, I would not be able to reach the dream of becoming a doctor. I also want to express my deep gratitude to them for having opened my mind to the most beautiful side of reality.

I want to thank first the two persons who have guided me in this adventure: Ángel Gómez and Alexandra Vázquez. I want to thank Alexandra specially for her patience, integrity, and good disposition to understand me and hear my complains and doubts, even when they were unfounded. Her equanimity has inspired me thorough all the way and drove me to try harder at the moments in which the task seemed too heavy and I felt the desire to give up and escape. I want to thank Ángel for his boundless passion and creativity, his equalitarianism and trust, and, above all, for his heart of gold. Seeing him genuinely worrying for the wellbeing and interests of his students, coworkers, collaborators, research participants, and society at whole has sparked the fire of authentic friendship within me and has given this adventure a meaning and significance much deeper and rich than what normally entails obtaining an academic degree. The dedication of Ángel and Alexandra has exceeded in much what is to be expected from PhD advisors; and their intellectual brilliance is what explains all the findings and discoveries that have been reached. They are the giants who put me on their shoulders so that I can see further away.

I want to thank Mercedes López and Marco Brambilla too. Mercedes was my master advisor and awoke the scientific curiosity in me, fed me with wise counsels and hope, and opened her home and soul to me, making me feel a warm in the chest that will be always there. Without her insistence, I would never have tried to obtain a PhD degree. Marco hosted me during a 3-month stay at the University of Milano-Bicocca and was the best imaginable host possible. He is extraordinarily intelligent and sociable and, what is more important, he is an extremely good person too.

I also feel gratitude towards the remaining members of Ángel's research team, Lucía López Rodríguez, Borja Paredes, Mercedes Martínez, Laura Blanco, Beatriz Alba, and Hend Bautista; sharing experiences and confidences with them has made the path much funnier and lighter.

The voluntary collaboration of thousands of participants from UNED and several dozens of imprisoned individuals has made this adventure possible. Talking to them and listening to their life stories has made me more conscious than ever of the complexity of reality and the importance of love, compassion, and social justice. Even though my gratitude will not reach most of them, I feel profoundly indebted.

I wish to thank Ramón Maiden too. He is a groundbreaking and unique artist who is always trying to capture what he contemplates looking deeply into the human spirit. The image of the keeper harboring a dragon on the chest that appears immediately after the cover of this thesis is his, and it is the most galvanizing representation of identity fusion that I was able to find.

Lastly, I want to thank my three mothers and my closest friends. They are the ones who have been always accompanying me and those who placed the pieces that constitute my being.

AGRADECIMIENTOS

En esta etapa de mi desarrollo académico y personal, no puedo evitar volver la mirada hacia atrás y observar que el camino que queda detrás es largo. Mayormente lo he recorrido como los exploradores de los países exóticos que admiraba de niña, llena de determinación y optimismo, y, como ellos, he encontrado algunos tesoros que me han hecho inconmensurablemente rica. El más valioso de todos no es otro que las personas que han caminado conmigo. Quiero dedicarles unas palabras para reconocer que, sin su ayuda, no sería capaz de alcanzar el sueño de convertirme en doctora. También quiero expresarles mi más profundo agradecimiento por haber abierto mi mente al lado más hermoso de la realidad.

Quiero darles las gracias primero a las dos personas que me han guiado en esta aventura: Ángel Gómez y Alexandra Vázquez. Quiero agradecer especialmente a Alexandra por su paciencia, integridad y buena disposición para entenderme y escuchar mis quejas y dudas, incluso cuando eran infundadas. Su ecuanimidad me ha inspirado profundamente durante todo el camino y me ha impulsado a esforzarme más en los momentos en los que la tarea parecía demasiado dura y sentía el deseo de darme por vencida y escapar. Quiero agradecer a Ángel por su pasión y creatividad ilimitadas, su igualitarismo y confianza y, sobre todo, por su corazón de oro. Verle preocuparse genuinamente por el bienestar y los intereses de sus estudiantes, compañeros de trabajo, colaboradores, los participantes en las investigaciones y la sociedad en su conjunto ha encendido el fuego de la auténtica amistad dentro de mí y le ha dado a esta aventura un sentido y significación mucho más profundos y ricos de lo que normalmente implica la obtención de un certificado académico. La dedicación de Ángel y Alexandra ha superado en mucho lo que cabe esperar de los directores de doctorado, y su brillantez intelectual es lo que explica todos los hallazgos y descubrimientos alcanzados. Ellos son los gigantes que me suben sobre sus hombros para que pueda ver más lejos.

Quiero darles las gracias también a Mercedes López y Marco Brambilla. Mercedes fue mi directora de master y despertó en mí la curiosidad científica, me alimentó con consejos sabios y esperanza, y me abrió su hogar y su alma, haciéndome sentir un calor en el pecho que siempre estará ahí. Sin su insistencia, nunca hubiera intentado obtener un doctorado. Marco me acogió durante una estancia de 3 meses en la Universidad de Milano-Bicocca y fue el mejor anfitrión imaginable. Es extraordinariamente inteligente y sociable y, lo que es más importante, es una persona extremadamente buena también.

También siento gratitud hacia los restantes miembros del equipo de investigación de Ángel, Lucía López Rodríguez, Borja Paredes, Mercedes Martínez, Laura Blanco, Beatriz Alba y Hend Bautista; compartir experiencias y confidencias con ellos ha hecho que el camino sea mucho más divertido y ligero.

La colaboración voluntaria de miles de participantes de la UNED y varias decenas de prisioneros ha hecho posible esta aventura. Hablar con ellos y escuchar sus historias me ha hecho ser más consciente que nunca de la complejidad de la realidad y de la importancia del amor, la compasión y la justicia social. Aunque mi gratitud no llegará a la mayoría de ellos, me siento profundamente endeudada.

También deseo darle las gracias a Ramón Maiden. Es un artista rompedor y único que siempre está tratando de capturar lo que contempla mirando en profundidad al espíritu humano. La imagen del guardián albergando un dragón en el pecho que aparece inmediatamente después de la portada de esta tesis es suya, y es la representación pictórica de la fusión de la identidad más estimulante que he podido encontrar.

Por último, quiero darles las gracias a mis tres madres y a mis amigos más cercanos. Ellos son los que siempre me han estado acompañando y los que han colocado las piezas que constituyen mi ser.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The psychological processes that drive individuals to engage in extreme personal sacrifices for the groups that are cherished to them are of great theoretical and practical relevance. Identity fusion, a feeling of oneness with a group so intense that dissipates the border that normally separates the personal self - the characteristics of individuals that make them unique - and the social self - the characteristics of individuals that derive from their group membership -, has emerged as one of the most potent predictors of extreme pro-group acts, including fighting, dying, and even killing for the group and each of its members. Although previous research suggests that morals could play an important role in the determination of the nature of identity fusion, its underlying mechanisms and its consequences, the relationships between identity fusion and the phenomena typically studied by moral psychologists have not been systematically addressed. With the present thesis, we aim to fill this gap in our knowledge tying identity fusion to several of the insights offered by moral psychology to try to find answers to some questions about the causal antecedents of identity fusion, the factors that regulate its consequences, and the motivational mechanisms that explain its effects.

In **Chapter 1**, we provide a brief overview of the identity fusion construct; explain what led us to assume that morality may be critical to reach a deep understanding of its nature, consequences, and associated mechanisms; and arrange the lines of research included in the dissertation. In **Chapter 2**, we conduct a literature review to summarize the chief theoretical and empirical findings about identity fusion that have been reached in recent years, offer an overview of the main misconceptions and untested assumptions about this construct, and spotlight some opportune hypotheses for future research. In **Chapter 3**, we explore whether admiration, an emotion elicited by the observation of laudable behaviours, causes a chain reaction augmenting identity fusion and ultimately leading individuals down a path toward behavioural extremism. In **Chapter 4**, we test whether moral beliefs about the justifiability of

violence moderate the relationship between identity fusion and proneness to engage in extreme violent and nonviolent pro-group actions. In **Chapter 5**, we examine if automatic, instinctive, and impulsive feelings of visceral responsibility based on care mediate the relation between identity fusion and the will to engage on extreme acts. Lastly, in **Chapter 6**, we conclude by summarizing the results, bringing up the limitations of the different lines of research, and pointing out some paths for a future research agenda.

RESUMEN

Los procesos psicológicos que impulsan a los individuos a realizar sacrificios extremos por los grupos que aprecian tienen una gran relevancia teórica y práctica. La fusión de la identidad, un sentimiento de unidad con un grupo tan intenso que disipa la frontera que normalmente divide el yo personal – las características de los individuos que los hacen únicos – y el yo social – las características de los individuos que derivan de la pertenencia grupal –, ha emergido como uno de los predictores más potentes de los actos pro-grupales extremos, incluyendo luchar, morir e incluso matar por el grupo y cada uno de sus miembros. Aunque la investigación previa sugiere que la moralidad podría desempeñar un papel importante en la determinación de la naturaleza de la fusión de la identidad, sus mecanismos subyacentes y sus consecuencias, las relaciones entre la fusión de la identidad y los fenómenos típicamente estudiados por los psicólogos morales no se han abordado sistemáticamente. Con la presente tesis, pretendemos subsanar esta laguna en nuestro conocimiento vinculando la fusión de la identidad con varias de las aportaciones de la psicología moral para tratar de encontrar respuestas a algunas cuestiones sobre sus antecedentes causales, los factores que regulan sus consecuencias y los mecanismos motivacionales que explican sus efectos.

En el Capítulo 1 presentamos un breve resumen del constructo de la fusión de la identidad, explicamos qué nos lleva a asumir que la moralidad podría ser crítica para alcanzar un conocimiento profundo de su naturaleza, consecuencias y mecanismos asociados, y exponemos las líneas de investigación incluidas en la disertación. En el Capítulo 2, realizamos una revisión de la literatura para resumir los principales hallazgos teóricos y empíricos sobre la fusión de la identidad que se han alcanzado durante los últimos años, ofrecer una descripción general de los malentendidos y suposiciones no probadas más importantes sobre este constructo y destacar algunas hipótesis oportunas para la investigación futura. En el Capítulo 3 exploramos si la admiración, una emoción desencadenada por la observación de

comportamientos loables, provoca una reacción en cadena que aumenta la fusión de la identidad y en ultima instancia conduce a los individuos por un camino que lleva al extremismo conductual. En el Capítulo 4, ponemos a prueba si las creencias morales sobre la legitimidad de la violencia moderan la relación entre la fusión de la identidad y la propensión a realizar acciones pro-grupales extremas de carácter violento y no violento. En el Capítulo 5, examinamos si la experiencia de sentimientos automáticos, instintivos e impulsivos de responsabilidad visceral basados en el cuidado median la relación entre la fusión de la identidad y la voluntad de participar en actos extremos. Por último, en el Capítulo 6, concluimos la disertación resumiendo los resultados, planteando las limitaciones de las distintas líneas de estudio y señalando algunas vías para la investigación futura.

CHAPTER 1

Theoretical Review

> You have noticed that the truth comes into this world with two faces. One is sad with suffering, and the other laughs; but *is the same face, laughing or weeping.*
>
> Black Elk (Neihardt, 2014; cursives are ours)

Black Elk, a holly man and warrior of the Oglala Lakota people, saw two antithetical faces in the same truth. Looking at the personal sacrifices that humans make for the groups that are dear to them invites us to conclude that we have two contrary sides too. Sometimes, people engage in marvellous deeds for such groups, like when they give their time, money, freedom, energy, and even their lives to protect the well-being of its members and guarantee their rights and security. However, individuals also perform terrible sacrifices to promote the interests of these groups, such as immolating themselves to kill innocent civilians, risking their lives in unnecessary wars, or violently retaliating for previous offences. As the two faces of truth about which Black Elk spoke, both of our sides can be conceived as two different manifestations of one underlying reality, because they often emerge from the same basic psychological state: identity fusion.

Identity fusion is a deep feeling of visceral oneness with a group characterized by a strong commitment to the group as a whole - collective ties - as well as to each of its members - relational ties - (Gómez et al., 2019; Swann et al., 2009, 2012). Conjointly, these attachments lead strongly fused individuals to react to the threats to the group as if they were personal threats (Gómez, Brooks et al., 2011; Swann et al., 2009) and trigger extraordinary feelings of personal agency that propel them to engage in great risks and personal sacrifices to protect the group and its members (Swann, Gómez, Huici et al., 2010). Since the inception of the construct, dozens of lab and field studies conducted by multidisciplinary teams of the five continents have shown that identity fusion is a robust predictor of extreme behaviours, including dying,

fighting, engaging in violence, performing costly sacrifices, and even killing for the group or its members (for reviews see Fredman et al., 2015; Gómez & Vázquez, 2015; Swann et al., 2012; Swann & Buhrmester, 2015; Whitehouse, 2018).

We chose to start the present thesis by conducting an extensive review of the theoretical and empirical contributions to identity fusion since 2015, the year in which the latest comprehensive reviews by one or both co-originators of the theory, William B. Swann Jr. and Ángel Gómez, were published (Fredman et al., 2015; Gómez & Vázquez, 2015; Swann & Buhrmester, 2015). Our first goal was to write a short manuscript for interested readers (1) providing an overview of the main findings reached since then; (2) addressing the doubts, misconceptions, and untested assumptions about identity fusion; and (3) pointing out some future avenues of research.

While conducting the review, we noticed that, despite researchers had discovered several points at which identity fusion intersects with the phenomena typically studied by moral psychologists (e.g., Carnes & Lickel, 2018; Heger & Gaertner, 2018; Kunst et al., 2018; Swann, Buhrmester et al., 2014; Swann, Gómez et al., 2014), there were still many open questions about the relationship between identity fusion and morality; and we decided to try to find answers to some of them. Thus, the second goal of this thesis was merging identity fusion research with the insights offered by research on morality to find responses to some of the most pressing questions posited by the existence of this feeling of visceral oneness. Specifically, moral philosophers and psychologists have been long talking about three moral processes that are relevant to reach a better understanding about identity fusion's antecedents, consequences, and associated mechanisms. Our efforts have been concentrated on the empirical study of these processes.

The first process is important to obtain a full picture of the causal precursors of identity fusion and is related to experiencing feelings of *admiration*, an emotion elicited by excellent

or virtuous behaviours that motivates the emulation of the ideals embodied by them (e.g., Immordino-Yang et al. 2009; Pohling & Diessner, 2016). The second process helps us disentangle the circumstances under which fused individuals' pro-group behaviour becomes negative or violent; and it is related to the heavier weight of proscriptive morality - rules and commitments to inhibit harmful acts - than of prescriptive morality - rules and commitments to engage in acts that help others - in regulating behaviour; and the potential deterrent impact of *moral beliefs against the use of violence* (Janoff-Bulman et al., 2009). Finally, the third process sheds some light into the mediating machineries that explain why strongly fused individuals are so astoundingly inclined to engage in extreme pro-group acts; and it is related to the experience of feelings of automatic or *visceral responsibility* described by some care ethicists (e.g., Noddings, 2003; Tronto, 1993, 1998).

Throughout this introductory chapter, we present the theoretical framework of the thesis. On the first section, we provide a short review of the identity fusion construct. On the second section, we offer a summary of the literature on admiration and explain in detail what drove us to assume that this emotion causes feelings of fusion. On the third section, we talk about the roles of proscriptive and prescriptive morality in motivating behaviour and explain why we supposed that moral beliefs regarding the justifiability of violence determine if strongly fused individuals manifest a violent side. On the fourth section, we outline some of the fundamental tenets of the ethics of care and clarify the reasons that led us to hypothesize that fused people are motivated to act on behalf of the target of fusion by feelings of visceral responsibility. Lastly, on the fifth section of the present chapter, we provide an overview of the thesis.

Identity Fusion

The construct of identity fusion was conceived around a decade ago to unravel why certain individuals, such as terrorists, social activists, or revolutionaries, are willing to display extreme pro-group acts (Swann et al., 2009).

Identity fusion is a visceral feeling of oneness with a group in which the personal self - the part of our identity related to our idiosyncratic characteristics - merges with the social self - the part of our identity related to our alignment to groups - without that none of both identities loses its integrity, which makes possible that they combine synergistically to motivate pro-group actions. Strongly fused individuals appreciate other group members because they belong to the group - collective ties - as well as because of their personal characteristics - relational ties - and establish with them psychological bonds like those that develop within the context of the biological family. Moreover, strongly fused individuals experience an intense feeling of personal agency - capacity to initiate and control intentional behaviours - that is used to selectively advance the interests of the group (Swann, Gómez, Huici et al., 2010). Lastly, because identity fusion gives meaning to the personal and social identity of strongly fused individuals, once people are fused, they tend to remain fused (see Gómez & Vázquez, 2015; Swann & Buhrmester, 2015; Swann et al., 2012).

Consequences of Identity Fusion

Since the inception of the theory, a great number of studies have shown that identity fusion is a remarkably robust predictor of willingness to engage in extreme pro-group behaviours. Among others, identity fusion predicts willingness to: (1) fight and die for the group and its members (e.g., Bortolini et al., 2018; Gómez, Brooks et al., 2011; Gómez, Morales et al., 2011; Heger & Gaertner, 2018; Swann et al., 2012); (2) volunteer for armed combat (Gómez et al., 2017; Kunst et al., 2018; Whitehouse et al., 2014); (3) make costly sacrifices for the values that are sacred for the group (Sheikh et al., 2016); (4) die to spare the lives of one or several ingroup members in different versions of the trolley dilemma (Gómez,

Brooks et al., 2011; Swann, Gómez et al., 2014; Swann, Gómez, Dovidio et al., 2010); (5) kill terrorists who threaten the group in an intergroup version of the trolley dilemma (Swann, Gómez, Dovidio et al., 2010); and (6) abandon valued personal relationships to belong to the group (Swann et al., 2015). In addition, identity fusion has been recently identified as a top risk factor predicting radical intentions in a meta-analysis including dozens of alternative predictors (Wolfowicz et al., 2021).

Other research has showed that fusion could predict current extreme behaviours and not only action intentions; and that some of such behaviours might even entail harming other people or turning to violence. For example, Whitehouse et al. (2014) interviewed a sample of Libyan revolutionaries opposing the regime of Gaddafi and discovered that frontline combatants were more fused with their own battalion than individuals who only provided logistical support and, accordingly, were less exposed to physical risks and suffered less casualties. Gómez et al. (2017) interviewed imprisoned ISIS terrorists and combatants in the frontline against ISIS and found that all of them were fused with their group. And Ferguson and McAuley (2020) performed a narrative and interpretative phenomenological analysis of the discourse of Northern Irish loyalists and republican paramilitaries and found that identity fusion played an important role in sustaining their extremism.

Additionally, other studies have found that fusion predicts non-extreme behaviours that are aimed at benefiting the group towards which fusion is experienced. These studies show, for example, that strongly fused people are more willing to write support notes to group members who have been victims of a terrorist attack (i.e., the Boston Marathon bombings; Buhrmester et al., 2014); remain in the group after having been ostracized (Gómez, Morales et al., 2011); deny the group's wrongdoing (Besta et al., 2014); and defend its reputation (Ashokkumar et al., 2019). Also, strongly fused individuals are more inclined to race a fusion-related avatar at high speed (Swann, Gómez, Huici et al., 2010); maximize the economic advantage of their

group over other groups at their own cost (Buhrmester et al., 2018); and make monetary donations to group members in a situation of need (Buhrmester et al., 2014; Swann, Gómez, Huici et al., 2010).

Although identity fusion was initially conceived as a particular type of bond with a group, recent research has extended the scope of the construct to cover the relationships that people establish with other entities, such as other individuals (e.g., Joo & Park, 2017; Kunst et al., 2019; Vázquez, Gómez, Ordoñana et al., 2017), animals (Buhrmester, Burnham et al., 2018), brands (Krishna & Kim, 2021), or values and convictions (Fredman et al., 2017; Martel et al., 2021). In all cases, fusion motivates people to engage in several behaviours to defend and protect the target with which they are fused.

Along with identity fusion with a group, the type of fusion that has received the most attention is interpersonal fusion, or fusion with a particular individual, such as the romantic partner, a sibling, or a twin. Results of this research show that identity fusion with an individual predicts willingness to engage in costly sacrifices (Vázquez, Gómez, Ordoñana et al., 2017) and fight and die for her or him (Joo & Park, 2015, Vázquez et al., 2015). Fusion with an individual also predicts, along other non-extreme things, more constructive ways to cope with conflicts (Walsh & Neff, 2018), and forgiveness after disappointments (Vázquez, Gómez, Ordoñana et al., 2017).

Causal Antecedents of Identity Fusion

Research on the causal antecedents of identity fusion has been mostly centred in two mechanisms signalled by Swann et al. (2012): shared biology and shared experiences. According to their view, identity fusion could have evolved during the tribal age to mark the boundaries of local groups of genetically related individuals and promote the inclusive fitness - the number of genes that an individual pass down to the next generation through its own offspring and that of genetical relatives - of members of such groups via self-sacrifices. They

posited too that sharing intense emotional experiences with other group members could cause fusion. Results of the studies conducted to date support both arguments.

Regarding *shared biology*, Vázquez et al. (2015) found that fusion with extended groups - groups so big that its members do not personally know each other - occurs when the ties that individuals establish with their siblings are projected to the group. It has been found too that monozygotic twins are more fused with their sibling than dizygotic twins, what indicates that the degree of genetic relatedness is an important causal factor (Vázquez, Gómez, Ordoñana et al., 2017). Swann, Buhrmester et al. (2014) discovered that, when participants of the five continents were asked to indicate the group for which they would die, most of them mentioned their family and were fused with it. Also consistently with the last finding, a field study with young men of the Mosul area (Iraq) short after the end of the territorial occupation by ISIS showed that fusion with family had a higher prevalence than fusion with close friends, one's tribe, Sunni Muslims, and all Muslims, respectively (Atran et al., 2018).

Concerning *shared experiences*, numerous studies have manifested that sharing intense emotional experiences with other group members augments identity fusion (e.g., Kapitani et al., 2019; Misch et al., 2018; Newson, et al., 2016; Tasuji et al., 2020). The effect of shared experiences is particularly strong when they are intense enough to make individuals reflect about their meaning (e.g., Buhrmester, Newson et al., 2018; Kavanagh et al., 2020; Newson et al., 2016); are negative, dysphoric, or traumatic (e.g., Jong et al., 2015; Segal et al., 2018; Whitehouse et al., 2017); or can be attributed to the will of supernatural agents (Segal et al., 2018). Also related to that, taking part in collective gatherings or rituals seems to cause and maintain fusion (e.g., Lobato et al., 2019; Lobato & Sainz, 2021; Páez et al., 2015; Zumeta et al., 2016).

Recently, scientists have begun to bring forward other potential causes of identity fusion too. For example, some authors have found that low cognitive flexibility (Zmigrod et

al., 2018, 2019) and perceived national continuity (Siromahov et al., 2020) could shape our personal view of national identity and promote identity fusion with national groups. Whitaker et al. (2021) used an agent-based modelling approach to study the evolutionary origins of identity fusion, assuming that individuals who are fused with a group are hypersensitive to perceived hypocritical group support from others. They discovered that all what was needed to promote the development of fusion within their simulation was simple self-referential judgment and ignorance of hypocrites. Kunst el al. (2018) noticed that political struggles (e.g., forced occupation of an outgroup's territory) may bolster fusion with groups to which people do not belong. And Carnes and Lickel (2018) observed that perceiving that the group shares core moral convictions also predicts fusion.

In addition to that, other researchers have started to study the ontogenetic development of identity fusion. Gaviria et al. (2015) conducted a study with school children (6-12 years old) and discovered that they were not able to experience a state of fusion like the one found among adults due to the lack of development of their personal identity. Given that children of this age have not built an image of themselves as stable, continuous, and singular beings, they are unable to establish the aspects of identity fusion related to the personal self, such as feelings of personal agency. Nevertheless, children can feel viscerally connected to a group and be willing to self-sacrifice for it. Gaviria et al. (2015) called the state of connection to the group that children experience "protofusion". Reese and Whitehouse (2021) further suggest that whereas feelings of visceral connection during early childhood may be triggered by phenotypic similarity, fusion based on episodic memories of shared experiences is not possible until mid-adolescence and may rely on suitable bonding experiences (e.g., rites of passage, emotionally intense team sports).

Factors that Moderate the Effects of Identity Fusion

In addition to revealing various antecedents of identity fusion, previous research has also identified several factors that attenuate or intensify its effects on pro-group behaviour. Paredes and collaborators, for instance, found that, among fused people, willingness to go to the extremes for the group diminishes when 1) they learn that most ingroup members would self-sacrifice for it due to their moral principles and emotions rather than to utilitarian reasons based on the number of lives that could be saved by their acts (Paredes et al., 2018); (2) have doubts about their own degree of fusion (Paredes et al., 2020); or (3) when the type of self-sacrifice demanded by the situation do not match their self-view (Paredes et al., 2021). Kunst et al. (2018) noticed that fused individuals are more willing to engage in costly acts for the group when they believe that they are morally obliged to do so than when they do not feel any obligation. And Swann, Buhrmester et al. (2014) observed that priming biological or symbolic characteristics shared by the group increases willingness to fight and die for it among the strongly fused.

Considering all the research focusing on the moderators of identity fusion, two findings are particularly noteworthy for the goals of the present thesis: (1) group threats maximize the effects of identity fusion (e.g., Buhrmester, Newson et al., 2018; Gómez, Brooks et al., 2011; Sheikh et al. 2016; Swann et al., 2009), but are not enough for the fused to engage in pro-group violence (e.g., Fredman et al., 2017); and (2) physical, psychological, and emotional activation also intensify its effects.

First, there are abundant empirical studies showing that *group threats* spur the strongly fused to defend and protect the group and its members with the same dedication with which non-fused individuals take care of themselves. For instance, Buhrmester, Newson et al. (2018) found that strongly fused people are more inclined to self-sacrifice for the group when they believe that its essence is being threatened. Sheikh et al. (2016) performed a study revealing that willingness to engage in costly sacrifices for a value was stronger among devoted actors -

participants who were fused with a group that shared this value and saw the value as sacred or non-negotiable - than among non-devoted actors only when they were reminded that the value was under threat. And Vázquez et al. (2020) conducted a series of studies using a videogame to measure aggressive inclinations and discovered that devoted actors only showed more aggressive inclinations that non-devoted actors when the idea that the outgroup represented a threat to the ingroup was accessible to them.

Together with the studies showing a strong degree of fusion among terrorists, fighters, football hooligans, and revolutionaries (e.g., Buhrmester et al., 2020; Gómez et al., 2017; Newson et al., 2018; Whitehouse et al., 2014), the former discovery can excite the idea that identity fusion always leads to an increase in willingness to engage in pro-group violence, but other studies showed that this is a misleading inference. For instance, Fredman et al. (2017) conducted a longitudinal study before, and soon after, the Palestinian Stabbing Intifada to test if fusion with Judaism and Israel predicted support for retaliatory policies under different levels of threat. They observed that fusion with Judaism was positively associated to support for retaliation; and that the relationship between both variables increased after the Intifada began. Nonetheless, this finding did not replicate for fusion with Israel, which, according to the authors, suggest that ideological factors might exert a considerable influence on the type of acts in which fused individuals engage. Specifically, that the ideologies communicated through religious rituals – e.g., the view that the members of the religious community are kindred spirits who are entitled to protect themselves against the threats posited by groups with different convictions - might be promoting or enabling hostile reactions towards outgroups.

Second, research has consistently shown that the effects of identity fusion on pro-group behaviour are especially strong under physical, psychological, or emotional activation. In respect to *physical activation*, Swann, Gómez, Huici et al. (2010) discovered that elevating participants' autonomic arousal through the practice of physical exercise - individually or in

group - amplified the disposition of strongly fused participants to engage in extreme pro-group acts; and Newson (2021) found that using psychostimulants (i.e., cocaine) may lead fused football fans to behave aggressively against their rivals. Regarding *psychological activation*, various studies have evidenced that questioning the personal or social identities of strongly fused individuals by giving them information that contradicts their self-views increases their willingness to engage in pro-group acts as compared to circumstances in which no inconsistent information is provided (Gómez, Brooks et al., 2011; Swann et al., 2009). Finally, concerning *emotional activation,* Swann et al. (2014) performed several studies with various adaptations of the trolley dilemma asking participants to choose between sacrificing their own life to save the lives of one or more ingroup members or letting several group members die to save themselves. They noticed that only participants who were strongly fused with the group preferentially chose sacrificing themselves; and that their inclination to self-sacrifice augmented when they had to decide under time pressure and their emotions were let to act unconstrained. Therefore, along with group threats, circumstances of physical, psychological, or emotional activation moderate the effects of identity fusion. Next, we summarize the results of research on the factors that explain its effects.

Factors that Mediate the Effects of Identity Fusion

Research on the explicative mechanisms of the effects of identity fusion has singled out various factors that mediate its effects. For example, Besta et al. (2018) discovered that, when people participate in mass gatherings, identity fusion promotes feelings of self-expansion - perceptions of new knowledge and an expanded perspective on reality - which in turn intensify group-efficacy beliefs and, ultimately, pro-group behaviour. Other authors have observed that the effect of fusion on pro-group behaviour is mediated by feelings of personal agency – capacity to initiate and control intentional behaviours - (Gómez, Brooks et al., 2011; Swann, Gómez, Huici et al., 2010). And Vázquez et al. (2019) conducted a study with siblings and

found that identity fusion predicted willingness to fight and die for the sibling trough the imagined personal consequences of losing him or her.

From all the identified mediating factors, three deserve a particularly high degree of attention in sight of our research: family ties, feelings of invulnerability, and emotional commitment.

First, according to identity fusion theory, the *family ties* that fused individuals establish with other group members should explain why they are so extraordinarily prone to act on behalf of the group. Supporting this idea, Swann, Buhrmester et al. (2014) showed that strongly fused individuals increased their perception of family-like ties with other group members and, in turn, augmented their willingness to engage in extreme actions on their behalf when they reflected on the core values of the group. And Buhrmester et al. (2014) demonstrated that family-like perceptions of fellow North Americans mediated the effect of identity fusion on support to the victims of the Boston Marathon bombings.

Second, and related to the previous reasoning, other researchers have found that *feelings of invulnerability* – the belief that nothing bad can happen to the self or the group - mediate the effect of fusion on pro-group behaviours, probably because the family-like ties that the fused experience led them to assume that the other members are also inclined to protect and defend the group (Gómez, Brooks et al., 2011).

Third, Swann, Gómez et al. (2014) wanted to know what strongly fused individuals have in mind when they decide to engage in self-sacrifices; and conducted a study with an adapted version of the trolley dilemma asking participants to record their thoughts while they were considering whether sacrifice themselves to save five ingroup members or let the group members perish to survive themselves. Results revealed that strongly fused individuals experienced intense feelings of *tension, anxiety, and emotional distress* in response to the plight

of the group members; and that these feelings mediated the positive relationship between identity fusion and self-sacrifice.

The research on the relationships between emotions and identity fusion is still scarce, but, in addition to the mediational role found by Swann et al., (2014), Kapitáni et al. (2019) conducted a study showing that the experience of positive affect during the inauguration of President Trump augmented identity fusion with the USA and people who shared participants' political views, which hints to the possibility that other-directed discrete positive emotions may promote fusion. We anticipated that one of such emotions, admiration, could be one of the antecedents of identity fusion.

Admiration and Identity Fusion

Admiration is an other-oriented emotion triggered by models of virtue or excellence (Imordino-Yang et al., 2009; Onu et al., 2016; Pohling & Diessner, 2016). Feeling admiration is usually described as a transformative experience that causes major personal changes, including self-improvement, the internalization and emulation of the ideals embodied by such models, and a focus on promoting the well-being of others and helping the group to achieve its ends (e.g., Algoe & Haidt, 2009; Haidt, 2003; Pohling & Diessner, 2016).

Research on the behavioural effects of admiration has revealed that admiring exemplars who engage in personal sacrifices for a cause positively relates to willingness to self-sacrifice for it (Bélanger et al., 2014). There is also an ample corpus of studies showing that admiration encourages prosocial acts, such as charitable donations, volunteering, or altruism (e.g., Cox et al., 2010; Schnall et al., 2010; Vianello et al., 2010, Vier & Abrams, 2010). Additional evidence shows that the effects of admiration extend to behaviours differing from those that triggered the emotion in the first place (van de Ven et al., 2018). And Sweetman et al. (2013) demonstrated that this emotion may even have the force needed to invigorate extreme personal risks and self-sacrifices; specifically, that Hong Kong citizens were more willing to take

political actions against the Chinese neo-authoritarian regime when they experienced admiration towards pro-democracy protesters than when they did not feel it. But why does admiration have these effects?

Theoretical explanations of the consequences of admiration are grounded in the social nature of human beings. Multilevel selection theory posits that admiration contributes to the "survival" of the group by motivating individuals to put their acts at its service (Pohling & Diesner, 2016). Socio functional approaches to the study of emotions state that admiration promotes greater group cohesion and commitment among its members (Stellar et al., 2015). And the most extended conception of admiration sees it as a catalyst for the formation of social bonds (e.g., Algoe & Haidt, 2009).

The thesis that admiration leads to feelings of connectedness to other individuals and groups has received some neurological and psychological support (e.g., Englander et al., 2012; Vianello et al., 2012), but the kinds of bonds that admiration causes have not been comprehensively researched. We supposed that admiration may fuse people to their group or convictions making them prone to engage in extreme self-sacrifices for them. Given that admiration is an emotion elicited by virtuous acts (e.g., Cox et al., 2010; Schnall et al., 2010; Vianello et al., 2010), we also thought that this emotion augments identity fusion by increasing the salience of the sacrifices that the target of admiration made for the group.

Importantly, the choices that individuals make about how they define their group determine whether such personal sacrifices translate into prosocial or aggressive behaviour; and, when violence is seen as a morally justified or necessary response to a violation against the group, people can even come to celebrate the most horrendous acts as manifestations of virtue (Reicher et al., 2008). In this moral universe, those who engage in violence are seen as bearers of moral strength, whereas those who doesn't are judged as morally suspect. Thus, admiration may lead people down a path of behavioural extremism and augment their

willingness to engage in violent and nonviolent radical acts. We supposed that another of the factors studied by moral psychologists, moral beliefs about the justifiability of violence, could regulate the type of extreme behaviours in which strongly fused individuals engage.

Beliefs in the Moral Justifiability of Violence and Identity Fusion

We have talked about the results of research showing that identity fusion does not necessarily motivates pro-group violence when the group is under a serious threat (Fredman et al., 2017), what indicates that ideological factors, such as moral beliefs, may be an important to figure out the type of pro-group behaviour (violent vs. non-violent) in which fused individuals are prone to engage. Two complementary lines of research also point to the fact that *beliefs in the moral justifiability of violence* may determine the consequences of fusion.

First, as previously explained, in their series of studies with the trolley dilemma, Swann, Buhrmester et al. (2014) noticed that fused individuals were more willing to sacrifice themselves for their fellow group members when reflective control was impeded and that their willingness to self-sacrifice was driven by emotional commitment to the group. Besides of that, the authors found that, despite all participants believed that self-sacrifice was the morally correct option, only fused participants preferentially chose sacrificing themselves and they were not sensitive to pragmatic arguments based on the number of lives that could be spared by their actions. Combined, these results indicate that the association between identity fusion and self-sacrifice may be due to a moral imperative that compels fused people to act as keepers of the group and "give up what is needed to preserve and promote the well-being of the group and its members" (for more information about type of moral imperative that might motivate strongly fused individuals; see Greene, 2013).

Second, Janoff-Bulman et al. (2009) presented evidence showing that there are two systems of moral regulation that differ in their motivational force: the prescriptive and the proscriptive regulation systems. The *prescriptive moral system* promotes behaviours aimed at

helping others by bolstering their well-being or alleviating their suffering, whereas the *proscriptive system* inhibits harmful behaviours. Given that bad events have a stronger psychological effect than good ones (e.g., Baumeister et al., 2001; Rozin et al., 2001), the avoidance focus on negative (vs. positive) outcomes of proscriptive morality leads to an asymmetry in motivational force that makes the proscriptive system stricter and more demanding.

Integrating the two previous lines of research helps us understand why identity fusion is not inherently related to violence as well as to predict when is more probable that fused individuals display a violent side. Concretely, strongly fused individuals feel morally prompted to promote the interests of the group and its members, which is a type of motivation regulated by the prescriptive system. Contrarily to that, moral beliefs against violence regulate behaviour through the proscriptive system. Given that this last system is more inflexible than the former one, we reasoned that strongly fused individuals are more willing to engage in violent behaviours for the group *only* when they think that violence is morally justifiable, but not when they hold beliefs that clearly oppose the use of violence. We also though that moral beliefs about violence should not influence strongly fused people's inclination to engage in extreme nonviolent pro-group acts.

Moral beliefs about violence may have a regulatory function among the strongly fused and determine the nature of the behaviours in which they choose to engage. We assumed that another moral factor, feelings of visceral responsibility, could explain why they are so prone to engage in extreme self-sacrifices.

Visceral Responsibility and Identity Fusion

Several results of identity fusion research, such as those showing that the relationship between fusion and pro-group behaviours is explained by family ties and emotional commitment (e.g., Swann, Buhrmester et al., 2014; Swann, Gómez et al., 2014), or that fused

individuals are more inclined to engage in extreme acts when they are physically, psychologically, or emotionally activated (e.g., Swann et al., 2009; Swann, Gómez et al., 2014; Swann, Gómez, Huici et al., 2010), indicate that this feeling of extraordinary connection may instigate people to engage in self-sacrifices for the target of fusion automatically, driven by feelings of *visceral responsibility* similar to those outlined by care ethicists.

Care ethics originated at the end of the 20th century as an alternative to the abstract and rationalistic view of morals adopted by deontologism and the liberal conception of justice (Gilligan, 1982), which are based on sheer adherence to duty and cannot accommodate behaviours that go beyond what can be rationally justified, such as those entailing critical personal sacrifices for others (e.g., Slote, 2007).

Departing from the previous view, care ethics conceives caring as grounded in feelings, receptivity to the needs of others, and responsiveness (e.g., Noddings, 2003; Tronto, 1993, 1998). Advocates of the ethics of care postulate that, when we engage with others in caring relationships, our passions give rise to automatic or impulsive feelings of responsibility for their well-being (e.g., Tronto, 1993, 1998), causing a motivational shift in which behaviour is directly determined by the needs of the cared for without any need for abstract reasoning or justification (Noddings, 2003). People who care for others are impelled to act as on their own behalf, but on behalf of the others; and, when that happens, "want" and "must" become the same thing (Noddings, 2003).

The type of feelings of visceral responsibility described by care ethicists can explain why people engage in prosocial behaviours exceeding the expectations established by duty; and we thought that experiencing them may be one of the mechanisms underlying the effect of fusion on extreme acts. Although the existence of this state can convey the impression that the strongly fused are particularly susceptible to repeatedly engage in costly behaviours for exploitative others, there are several mechanisms that could preclude this. For instance, care

ethicists have asserted that trust and solidarity are a requirement of care (e.g., Tronto, 2013; van Nistelrooj, 2015), what intimates that a shared sense of responsibility is needed to maintain a high level of self-sacrifice.

Overview of the Thesis

The present thesis had two chief objectives. Our first goal was to analyse the principal misunderstandings and the main doubts and untested assumptions that have arisen during the years in light of the original postulates of identity fusion theory after conducting a previous extensive review of the most recent literature. Our second goal was to merge identity fusion research and moral psychology by examining if (1) admiration "fuses" individuals to groups and causes; (2) prescriptive beliefs about violence regulate the type of extreme behaviour (violent vs. non-violent) in which the strongly fused intend to engage; and (3) strongly fused individuals engage in extreme prosocial behaviours because they feel visceral responsibility towards the fusion target.

Chapter 2 is titled "Recent advances, misconceptions, untested assumptions, and future research agenda for identity fusion theory" and is a narrative review article. It presents an updated overview of the theoretical and empirical advances in our understanding of the nature, causes, consequences, underlying mechanisms, and applications of identity fusion since 2015 and up to 2020. And what is more important, it also addresses the main misunderstandings, doubts, and untested assumptions about the construct; and points out some promising avenues for future research.

Chapter 3 is titled "Admiration for Islamists encourages self-sacrifice through identity fusion" and is an empirical article. This article includes one quasi-experimental study with imprisoned Jihadist terrorists (Study 1) and five complementary experimental studies (Studies 2-6) showing that admiration causes identity fusion with group or ideological convictions and that, through this process, augments willingness to engage in extreme sacrifices. The studies

also reveal that admiration promotes identity fusion by rendering the sacrifices that the target of admiration made for the group cognitively salient (Studies 2-6); and clarify some of the boundary conditions under which admiration exerts its effects (Studies 2-5). The last experiment (Study 6) shows that the effects of admiration are long-lasting and are not restricted to behavioural intentions but extend to real prosocial behaviours.

Chapter 4 is titled "Identity fusion predicts violent behaviour when it is morally justifiable" and is also an empirical article. It includes two cross-sectional studies showing that moral beliefs regarding the justifiability of violence moderate the relationship between fusion and willingness to engage in violent and nonviolent pro-group acts. Concretely, the studies reveal that strongly fused individuals are more willing to act violently than weakly fused people if they believe that violence is morally justifiable, but not if they hold thoughts that clearly proscribe the use of violence. The studies also show that moral beliefs about violence do not influence strongly fused individuals' willingness to engage in nonviolent acts.

Chapter 5 is titled "Strongly fused individuals feel viscerally responsible to self-sacrifice" and is an empirical article too. This article presents two cross-sectional studies and one experiment manifesting that strongly fused persons are willing to engage in extreme behaviours on behalf of the target of fusion because they experience visceral responsibility towards it. They also reveal that perceiving that feelings of visceral responsibility are not shared mitigates the effects of fusion on extreme prosocial acts.

Lastly, Chapter 6 includes an overview of the practical and theoretical contributions of the thesis to identity fusion and moral psychology. It also highlights its main shortcomings and points out some ideas and hypothesis for future research.

All of the articles resulting from the current thesis have been published in scientific journals (see Annex A). We include the references for the publications as well as information on the impact factor and classification of the journals within the corresponding chapters.

CHAPTER 2

Recent Advances, Misconceptions, Untested Assumptions, and Future Research Agenda for Identity Fusion Theory

Reference

Gómez, Á., **Chinchilla, J.**, Vázquez, A., López-Rodríguez, L., Paredes, B., & Martínez, M. (2020). Recent advances, misconceptions, untested assumptions, and future research agenda for identity fusion theory. *Social and Personality Psychology Compass, 14*(6), e12531. https://doi.org/10.1111/spc3.12531. JIF, 2020: 2.52 (Q3).

Abstract

Just a decade ago, two psychologists, W. B. Swann Jr., and Á. Gómez, developed a new theoretical framework to explain extreme pro-group behaviours: *identity fusion theory*. *Identity fusion* refers to a visceral feeling of oneness with a group that motivates individuals to do extraordinary self-sacrifices on behalf of the group or each of its members. Since the formulation of the theory, interdisciplinary researchers of the five continents have conducted dozens of studies on identity fusion, both in laboratory and field settings. Research has deepened into the causes, consequences, underlying mechanisms, and applications of identity fusion. The development of fusion-based research has been steadfast and very prolific. Hence, the first section of the current manuscript includes an updated overview of this fast-growing literature. This increasing interest for the theory has however been accompanied by a series of misconceptions and untested research assumptions, which we address in the second and third sections of the paper, concluding with a final section suggesting a future research agenda. Our aim is to help those interested in knowing more about identity fusion or about the causal mechanisms that lead individuals to risk their life and personal wellbeing for a group discarding common misconceptions as well as formulating more precise and nuanced hypotheses for future research.

Keywords: Identity fusion, extreme behaviour, self-sacrifice, pro-group behaviour, misconceptions, oneness, social identity

Identity fusion was initially conceived as a visceral feeling of oneness with a group that predicts extreme pro-group behaviour with great fidelity. Strongly fused individuals experience an extraordinary sense of personal agency and reciprocal strength from group membership. Although identity fusion was originally conceptualized as a particular alignment of an individual with a group, the expansion of the theory has extended this mechanism to the relation that an individual can also develop to another individual, an animal, an object, or an activity. Fusion with any of these entities has implications for the way individuals behave as a result of being fused.

The original idea of *identity fusion* emerged in the aftermath of the 9/11 terrorist attacks in New York, and the Madrid Train Bombings. Two social psychologists, William B. Swann Jr. and Ángel Gómez, jointly conceived the concept of *identity fusion* as an attempt to explain why some individuals are willing to display extreme pro-group behaviours. Although the authors formally presented preliminary findings in 2005 and 2007[1], it was a decade ago when the first empirical publication appeared (Swann et al., 2009).

Identity Fusion Theory before 2015

After the emergence of the theory, dozens of experiments in the five continents conducted by multidisciplinary teams of researchers have shown that identity fusion is a reliable predictor of willingness to fight, kill, and die for one's group. Research conducted in the first five-years period after the formulation of the theory was typically focused on two main objectives. The first aim was finding empirical support for the four principles of the theory

[1] William B. Swann Jr. and Ángel Gómez presented identity fusion at the Preconference for the Society for Personality and Social Psychology Conference celebrated in New Orleans, 2005; the Social Psychology Winter Conference celebrated in Utah, 2005; and the Preconference for The Self and Identity celebrated in Chicago, 2007.

(Swann et al., 2012): the agentic-personal-self principle - fused individuals display high levels of personal agency that serves the group's agenda -, the identity synergy principle - personal and social identity combine synergistically to motivate pro-group behaviour -, the relational ties principle - appreciation of the unique personal identity of each ingroup member -, and irrevocability - once fused, tendency to remain fused -. The second goal was disentangling theoretical and empirically fusion from related but different socio-psychological constructs, especially from group identification (Tajfel & Turner, 1979).

These initial attempts to find empirical foundations for the theory as well as to articulate the key overlaps and differences between identity fusion and group identification produced several reviews mainly focused on how the findings supported each of the four principles (Buhrmester & Swann, 2015; Fredman et al., 2015; Gómez & Vázquez, 2015; Swann & Buhrmester, 2015, see Table 1 for a summary).

However, trusting that the theory was relatively well founded, these reviews also motivated researchers to search for answers to new, unexplored theoretical and empirical questions, and to extend the research to field studies with populations of interest to explore extreme behaviours, which marked a turning point on identity fusion research.

The theoretical and empirical contributions after 2015, have grown exponentially. Researchers from several disciplines (e.g., psychology, anthropology, psychobiology, neuroscience, political sciences, etc.), and from the five continents, have conducted laboratory and field studies to deepen into the nature, causes, consequences, underlying mechanisms, and applications of identity fusion theory. Nevertheless, as a consequence, this productive development has been accompanied by some misconceptions and untested assumptions. The present manuscript presents an overview of these theoretical and empirical contributions, addresses the main misconceptions and untested research assumptions, and concludes with some suggestions to establish a future research agenda.

Table 1. Key main differences and similarities between identity fusion and classic perspective on group identification

	Differences	
	Identity Fusion Theory	**Classic Social Identity Theory**
Focus of the theory	Extreme pro-group behaviour, enduring commitment, emphasis on intragroup relations	Intergroup relations
Relation between personal and social identities	Highly permeable	Lowly permeable
Salience of personal and social identities	Independent	Antagonistic
Role of personal identity	Highly influential	Weakly influential
Stability	High (Irrevocability)	Low (Context-dependency)
Type of ties involved	Relational and collective ties	Collective ties
Perception of ingroup members	Personalized (uniqueness-based)	Depersonalized (prototypical)
Group-related reasoning	Deontological	Utilitarian
Goals	Group and individual-serving	Group-serving
Predictive capacity of extreme behaviour	High	Moderate
Scales/questionnaires to be measured	High consensus	Moderate/low consensus
	Similarities	
	Identity Fusion Theory	**Classic Social Identity Theory**
Applications of the theory	Intragroup and intergroup processes	
Antecedents/Causes	Shared negative experiences (discrimination, suffering, natural disasters)	
Consequences for the self	Cognitive (representation of the self and the group, sensitivity to group-related information) Affective (self-esteem, group-related emotions)	
Consequences for the group	Group commitment, group cohesion, group loyalty, pro-group behaviour, etc.	

Main Advances and Discoveries since 2015 for Identity Fusion Theory

After 2015 identity fusion research experienced a shift. Believing that principles of the theory were quite supported, academics began to pose new challenges for extending the scope of the theory. At this time, they started to study processes like the development of identity fusion, its temporal stability, the possibilities of de-fusion, or fusion with entities different from the group to which individuals belong. Additionally, they tested novel hypotheses about the antecedents, consequences, and mechanisms of identity fusion. Moreover, methodological advances were made as well, and investigations were conducted with samples from special populations of interest. In this section, we present a summary of the main advances and discoveries of the last five years.

Ontogenetic Development

Gaviria, Ferreira, Martínez, and Whitehouse (2015) explored the developmental origins of fusion in school children - 6 to 12 years old -. Results indicated that children are not able to experience a state of fusion like that found in adults, due to the lack of development of their personal identity. Because children at this age have not developed an image of themselves as stable, continuous, and as singular individuals, they do not have established the aspects of identity fusion related to the self, as the feelings of personal agency or the belief that the self makes the group strong. Nonetheless, children can feel deeply connected with a group and express willingness to make sacrifices for it. They dubbed the state of connection to the group that children experience as “protofusion”.

Temporal Stability

Vázquez, Gómez, and Swann (2017) examined whether internal historic events (e.g., corruption scandals) threatening the group affect identity fusion. They found that average fusion scores declined after the occurrence of this type of events. However, this decline was

restricted to sentiments towards the group category - collective ties -, and it did not affect sentiments toward individual group members - relational ties -, or willingness to fight and die for the group, indicating that some aspects of identity fusion are more resistant to change than others.

De-fusion

Although one of the principles of identity fusion is its irrevocability, researchers have striven for reducing fusion with the group and/or its consequences. One of the main interests is the application of techniques that could reduce violent radicalization, at least when it is caused by fusion with the group. To that end, Gómez et al. (2019) conducted a series of experiments to find out if degrading either collective ties (i.e., sentiments toward the group as a whole) or relational ties (i.e., sentiments toward individual group members) lowered identity fusion and pro-group behaviour. Results showed that degrading relational as well as collective ties diminished fusion with the group and pro-group actions. On the other hand, although degrading collective ties reduced overall group identification, degrading relational ties only reduced scores on a single component of a multidimensional measure of group identification: ingroup solidarity (Leach et al., 2008).

Entities to be Fused with

One of the most fundamental innovations of this period is the notion that it is possible to be fused with different types of entities. In addition to the traditional conception of identity fusion, as the relation that an individual develops with a group (the ingroup, but also an outgroup that, for example, is oppressed, see Kunst et al., 2018), people can fuse *with another individual* as his/her romantic partner (Joo & Park, 2017; Walsh & Neff, 2018), his/her sibling (Vázquez, Gómez, Ordoñana et al., 2017), or a political leader (Kunst et al, 2019). But

individuals can also fuse *with an animal* (Buhrmester, Burnham et al., 2018), *a trademark* (Hawkins, 2019), or *a value or conviction*, as religion (Fredman et al., 2017).

The possibility that people can fuse with different entities represents a substantial departure from the original formulation of identity fusion and raises new questions that have not yet been resolved. For instance, the mechanisms that underlie fusion with a group and explain its consequences may be different from the processes involved in fusion with another type of entity. We will address this issue in the future lines of research section of the current manuscript.

Antecedents

Research on the causes of fusion has been extraordinarily prolific in this period. Studies have been centred in two causal factors proposed by Swann et al. (2012): shared biology and shared experiences. According to these authors, identity fusion could have evolved in tribal groups as a mechanism to demarcate local groups of genetically related persons, and to maximize the inclusive fitness - the ability of an individual organism to pass on its genes to the next generation including the shared genes passed on by close relatives - of individuals within these groups through the promotion of self-sacrificial behaviours. From these local groups, fusion could then be projected to extended groups - large groups of genetically unrelated individuals - as a result of perceptions of shared essence. Another mechanism that could explain the emergence of fusion is sharing emotionally intense, transformative experiences with other group members. Concerning shared biology, Vázquez, Gómez, Ordoñana et al. (2017) found that, as compared to dizygotic twins, monozygotic twins were more fused with their sibling. As predicted by Swann et al. (2012), twins' degree of genetical relatedness was positively related to identity fusion. Consistent with this hypothesis, a study conducted with young men of the Mosul area (Iraq) revealed that fusion with family was more prevalent than fusion with close friends, Muslims generally, Sunni Arabs, or one's tribe (Atran et al., 2018).

Whitehouse and associates have done an impressive amount of studies dealing with shared experiences as a potential cause of fusion. These studies demonstrate that sharing intense experiences with other ingroup members fosters fusion with the group (Kapitány et al., 2019; Misch et al., 2018; Newson et al., 2017), particularly if these experiences are negative, traumatic, or dysphoric (Jong et al., 2015; Segal et al., 2018; Whitehouse et al., 2017). Other studies also found support for the idea that participating in collective gatherings, like folkloric marches and religious celebrations, could increase fusion (e.g., Páez et al., 2015; Zumeta et al., 2016). Finally, recent investigations show that engaging in ritual practices and recalling episodic memories of a pilgrim's route contribute to maintain identity fusion (Lobato & Sáinz, 2019).

Several researchers have found other potential causes of fusion. Kunst et al. (2018) showed that political struggles (e.g., oppressive occupation of the outgroup) that clash with people's political beliefs might lead to fusion with groups to which individuals do not belong. Carnes and Lickel's (2018) manifest that perceiving that the group shares core moral beliefs or convictions can also cause fusion. Finally, Zmigrod, Rentfrow, and Robbins (2018, 2019) found evidence that cognitive inflexibility and ideological orientations could shape our personal sense of nationalistic identity, bolstering fusion with national groups.

Consequences of Identity Fusion

All of the studies conducted since 2015 confirm that fusion motivates individuals to engage in several kinds of actions to protect or defend the entity with which they are fused, in addition to the traditional outcome measure of willingness to fight and die for the group. What they do seems to be determined by the kind of entity they are fused with as well as by some situational factors (e.g., Fredman et al., 2017; Misch et al., 2018; Newson et al., 2016; Vázquez et al., 2017). For instance, people who strongly fuse with an outgroup that is victim of unjust treatment, like the Palestinians or the Kurds, are more willing to participate in extreme forms

of protest on behalf of the group (Kunst et al., 2018); whereas people who fuse with an activity related to a brand are more willing to spread negative world-of-mouth, to boycott the brand, and to avoid repurchasing the brand after a market disruption or a brand transgression (Hawkins, 2019). In contexts where the group's essence is threatened, strongly fused persons are especially likely to maximize the ingroup's advantage over the outgroup even at the expense of personal costs (Buhrmester, Newson et al., 2018). Other studies show that fusion with a group may engender lifelong loyalty to it (Newson et al., 2016), and that fusion with religion is positively related to the desire of retaliation after a threat to the religious group (Fredman et al., 2017). Importantly, it appears that highly fused persons are not only more willing to self-sacrifice for the group, but also to sacrifice the group and its members for their personal gain (Heger & Gaertner, 2018).

Identity fusion could also have some impact on our moral and socio-political preferences and well-being. For instance, Talaifar and Swann (2018) demonstrated that fusion with the country might break the political divide between liberals and conservatives with regard to their endorsement of the moral foundations of loyalty/betrayal, authority/subversion, and purity/degradation. Kunst et al. (2019) found that fusion with certain political leaders (i.e., Donald Trump) predicts willingness to endorse and engage in political violence (e.g., persecuting immigrants and political opponents). Recent work by Ashokkumar, Galaif, and Swann (2019), show that after a public transgression of the group, strongly fused individuals strive to protect the group's reputation. And Talaifar et al. (2020) found that students who were strongly fused with their university were more likely to remain in school up to a year later. Alternative studies indicate that one of the positive consequences of being fused with a group could be self-expansion (Besta et al., 2018), and that being fused with the romantic partner could lead to more constructive ways of coping with relationship conflicts and reduced vigilance for relationship threats (Walsh & Neff, 2018).

Underlying Mechanisms Related to Identity Fusion

Research has also tried to expand the factors that moderate or mediate the effects of fusion. With respect to the moderators, it has been discovered that strongly fused individuals are particularly willing to sacrifice for the group when the essence of the group is threatened (Buhrmester, Newson et al., 2018). Fused individuals are also more willing to go to the extremes for the group when they feel morally compelled to do so than when they do not experience any feelings of moral obligation (Kunst et al., 2018). In contrast, strongly fused individuals reduce their willingness to fight and die for their group when they learn that other ingroup members would self-sacrifice for it due to their moral principles and emotions toward the group as opposed to a pragmatic calculus about the costs and benefits associated to self-sacrificial behaviour (Paredes et al., 2018). Regarding the mediators, a recent research has revealed two additional mechanisms, feelings of self-expansion and group-efficacy beliefs, that operate sequentially (Besta et al., 2018). These studies were conducted during various mass gatherings, including music festivals, a demonstration of bicycle activists and cycling lovers, and a protest of LGBT right supporters. Results showed that, when people participate in crowd gatherings, identity fusion increases feelings of self-expansion resulting in new knowledge and an amplified perspective on reality, which in turn promotes group-efficacy beliefs and, ultimately, pro-group behaviour. Research on the mechanisms related to fusion has also examined the processes that amplify the effect of shared experiences on fusion. Several studies point out that shared experiences are especially effective to increase identity fusion when they are attributed to the will of a supernatural agent (Segal et al., 2018) and when they are intense enough to make individuals reflect about their meaning and believe that they have been personally shaped by them (Buhrmester, Newson et al., 2018; Newson et al., 2016).

Neural Bases

A series of investigations have used fMRI techniques to examine the neural bases of identity fusion and its correlates. Some authors have found that the level of fusion modulates the differential activation of the ventromedial portions of the prefrontal cortex (VMPFC) in response to fair (vs. unfair) money offers received from ingroup (vs. outgroup) members, suggesting that the activation of these portions of the brain may mediate the influence of fusion on our reaction to the behaviour of other individuals (Apps et al. 2018). Some others have tried to disentangle the neural correlates of the relation between identity fusion, sacred values, the will to fight, and violent extremism (Hamid et al., 2019; Pretus et al., 2018, 2019).

Methodological Advances

Jiménez et al. (2016) developed the Dynamic Identity Fusion Index (DIFI), which is a continuous adaptation of the pictorial item that was used in the seminal paper about identity fusion (Swann et al., 2009). The DIFI combines the simplicity afforded by a single pictorial item with the precision of a continuous measure, and it can be used off-line and on-line in traditional computers and touch-pad devices. It is particularly useful when researchers want to conceal that they are measuring fusion, or when they work with illiterate populations.

As we have showed, during the last five years the development of fusion-based research has been steadfast. Lab research has been systematically complemented with field studies conducted with special samples, like hooligans, twins, college fraternity/sorority members, military veterans, political partisans, martial arts practitioners, fighters against the Islamic State, or terrorists (e.g., Gómez et al., 2017; Kapitány et al., 2019; Newson et al. 2016; Whitehouse et al., 2018). There is even a behavioural economic experiment that contrasts the effect of fusion on eight different sociocultural groups ranging from foragers and horticulturalists to fully market-integrated individuals (Purzycki & Lang, 2019). Nonetheless, this increasing interest for the theory has come along with several misconceptions and untested

research ideas. We will offer some details about their merits and flaws in the following two sections.

Misconceptions about Identity Fusion

Scholars specialized in the study of radical behaviour and group processes have started to pay close attention to identity fusion. Without wishing to belittle their important contributions to the fusion research, we would like to clarify several misconceptions that can be grouped in two general categories: regarding the nature of identity fusion, and regarding its antecedents and consequences.

Misconceptions Regarding the Nature of Identity Fusion

Here we describe what we consider are the three most important misconceptions regarding identity fusion theory: its assimilation or subsumption to social identity theory, to communal sharing, or to a personal predisposition.

The most common inaccuracy is to assume that fusion and group identification are one and the same thing. Babinska and Bilewick (2018) presume that being fused with an extended group is identical to being identified with it, and Milla, Putra, and Umam (2019) seem to use both terms indistinctly. Vignoles (2017) suggests that fusion is subsumed within the conceptual sphere of identification, whereas some characteristics that are conventionally considered as the essential qualities of group identification (e.g., collective self-esteem, ingroup homogeneity) are mere correlates of it. Identity fusion, he adds, is a core dimension of group identification and should not be treated as a separate construct. These misconceptions are easily understandable because insofar as fusion and identification refer to the psychological ties that bind individuals to groups, both constructs are intimately related. However, unlike social identity theory (Tajfel & Turner, 1979) which is mainly concerned with intergroup relations and collective ties, identify fusion theory emphasizes the intragroup dynamics that prompt

individuals to sacrifice for their ingroup. Strongly fused individuals value the unique characteristics of fellow group members and develop family-like ties with them even when they are not personally acquainted with them (Swann, Gómez et al., 2014). Both fusion and social identity theories recognize the importance of collective ties, but the fusion approach specifically acknowledges the motivational role of the personal self and relational ties in predicting pro-group behaviour (see Gómez et al., 2019). Given these differences, it is not surprising that fusion and group identification are associated to different variables or differ in regard to their predictive power, as several studies indicate. Bortolini, Newson, Natividade, Vázquez, and Gómez (2018), for instance, measured fusion and identification (in this case with multidimensional and unidimensional scales) with three different groups to compare their effects on pro-group behaviour. Their studies found that fusion explained further variance than each of the identification measures. Recently, Gómez et al. (2019) found that undermining relational ties to fellow group members affected fusion, but not identification with the group. These and other results (e.g., Gómez, Morales et al., 2011; Swann, Gómez et al., 2014) support that "fusion" and "identification" should be treated separately.

Other authors consider identity fusion as one of the four basic relational modes: communal sharing (Fiske & Rai, 2015; Thomsen & Fiske, 2018). They take as proof of this equivalence some of the similarities that exist between the two constructs, for example, that both presuppose the experience of feelings of oneness with the group, are associated to caring for others, and may lead to "virtuous violence". Beyond these shared aspects that Fiske pertinently points out, there are remarkable differences between identity fusion and communal sharing that impede the matching of both constructs. Whereas in communal sharing relationships individual identities are disregarded and people are treated as equivalent and undifferentiated (Fiske, 1992), strongly fused persons recognize the unique personal identities as well as social identities of fellow group members (Swann et al., 2012). Thus, identity fusion

and communal sharing are based on distinct essential mechanisms and the former is more specific than the latter.

A third misconception is the assimilation of identity fusion to a personal disposition, akin to a personality trait, that leads individuals to fuse with groups. In opposition to that, research shows that fusion is the result of intragroup processes (e.g., sharing experiences, Jong et al., 2015; Segal et al., 2018; Whitehouse et al., 2017), that some of it components diminish when the group is threatened by internal events (Gómez et al., 2019; Vázquez et al., 2017), and that it does not correlate with a plethora of personality traits as the categories of personality of the Big Five (Gómez & Vázquez, 2015). Even though the former studies do not entirely preclude that identity fusion may be associated to some personal disposition not studied to date, they point to the idea that it is not a personality trait. We think of it as a stable and long-lasting psychological state, because the relational bonds to other group members that strongly fused people develop lock them into self-perpetuating interpersonal cycles (Swann et al., 2012).

Misconceptions Regarding the Antecedents and Consequences of Identity Fusion

In this subsection, we summarize the four more relevant misconceptions regarding the theory, as considering identity fusion as the unique or most potent predictor of extreme behaviour, criticizing the theory based on the interpretations that some have done of the original theory, believing that fused individuals cannot sacrifice for distant groups, or treating fusion as a transitory mechanism.

First, some authors appear to presuppose that identity fusion is the single or the most powerful predictor of self-sacrificial behaviour (Whitehouse, 2018). We do not share this point of view since human behaviour is complex and motivated by a myriad of factors. In interaction with other variables (e.g., intergroup threat), identity fusion can partly explain many of the sacrifices that individuals make for their groups. However, people do not always sacrifice themselves for their group; sometimes they do it to defend their ideals and sacred values

(Gómez et al., 2017). Hence, identity fusion will be especially predictive when self-sacrificial behaviour is group-oriented, but other variables can outperform fusion in predicting self-sacrificial behaviour if the goal is to protect one's values (Gómez et al., 2017). On the other hand, ideals and values could be causal antecedents of fusion (Carnes & Lickel, 2018; Kunst et al., 2018) or they can interact with fusion to stimulate pro-group behaviour as the devoted actor model states (Gómez et al., 2017; Sheikh et al., 2016; Vázquez et al., 2020).

Second, some authors support their criticisms on what some others interpret from identity fusion theory rather than on the empirical findings or the assumptions of the original authors. For example, Olivola (2018) asserts that many extreme forms of self-sacrifice occur without the participation of fusion. Similarly, Kiper and Sosis (2018) say that a theory about extreme self-sacrifice cannot be limited to identity fusion and group threats. We agree with the idea that identity fusion is not the single or most important cause of *any* extreme behaviour. The misunderstanding here resides on the Whitehouse's (2018) assumption that extreme forms of self-sacrifice for the group require a radical form of identity fusion with one's group.

Third, Crimston and Hornsey (2018) suggest that identity fusion cannot motivate sacrifices on behalf of distant groups to which individuals do not belong, like animals or disadvantaged outgroups, whereas moral expansiveness can explain them. Recent studies question this idea and demonstrate that identity fusion does stimulate extreme behaviour for distant groups too (Buhrmester, Burnham et al., 2018; Kunst et al., 2018). In particular, Kunst et al. (2018) found that people can fuse with oppressed outgroups and, in turn, engage in costly solidarity actions. Caring and being morally concerned for the wellbeing of the group or its members are core components of fusion (Fredman et al., 2015; Swann, Gómez et al., 2014). Hence, it is quite probable that fused individuals' moral realm is expanded to incorporate the groups with which they are fused.

And fourth, Wiessner (2018) assumes that fusion could be a fleeting state, and that people move in and out of the state of fusion to evaluate the personal risks and benefits associated with self-sacrifice. Contrarily to what Wiessner assumes, the irrevocability principle of fusion theory states that once fused, people will tend to remain fused (Swann et al., 2012), and previous research already showed that some components of fusion are heavily resistant to change (e.g., Vázquez et al., 2017).

Untested Assumptions About Identity Fusion

Some assume that identity fusion is uniquely associated with violence and negative behaviours. Even though lots of studies show that strongly fused individuals are more inclined to engage in violence on behalf of the group than weakly fused individuals (see Whitehouse, 2018), identity fusion can also promote prosocial behaviour such as providing various forms of support to the victims of the Boston Marathon bombings (Buhrmester et al., 2015), donating funds to ingroup members in need (Gómez, Morales et al., 2011), and donating time and money to the community after a natural disaster (Segal et al., 2018). Hence, fused individuals are more willing to do *intense sacrifices* for the group, but the specific manifestations (violent or prosocial) of this higher propensity to go to the extremes depend on ideological and contextual factors. Those persons who are fused with peaceful groups and have internalized the moral principle of not harming should not be distinctively inclined towards violence.

Finally, Lankford (2018) believes that fusion only relates to reported willingness to die for the group, but not to real willingness to die for it. It is difficult to address this fair criticism empirically, since we cannot ask individuals who already gave their lives for a group if they were fused. Even if we question individuals who are going to put their lives at risk, there is always place for harbouring a reasonable doubt about what they say. However, several investigations suggest that fusion predicts extreme behaviour and not only intentions. For

instance, a study conducted with Libyan revolutionaries (Whitehouse et al., 2014) showed that frontline combatants were more likely to be fused with their battalion than those who only provided logistical support and, in turn, were less exposed to death. Also, in a sample of transsexuals, those who were fused with their cross-gender group were more than twice as likely to have undergone irreversible surgical change of their primary sexual characteristics than non-fused participants two years after the assessment of fusion (Swann et al., 2015). And in interviews with imprisoned ISIS members and combatants in the frontline against ISIS, Gómez et al. (2017) found that all were fused with their group.

Future Lines of Research

As we have seen, our knowledge about identity fusion has steadily increased after Fredman et al.'s (2015) work. However, the more we test the theory, the more avenues are open to continue increasing our knowledge about it. Here we will summarize some future directions that identity fusion research could follow to make new advances, but also to deal with the misconceptions and untested assumptions that we have described.

Regarding how to reinforce the advance of the theory, we highlight the need to continue exploring the ontogenetic development of fusion. Research has showed that children up to 12 experience some kind of protofusion (Gaviria et al., 2015). The first empirical publications in the field (Swann et al., 2009; 2010) included high school participants between 15 and 16 years old. Then, there is a period between 12 and 15 years old that should be scrutinized to understand how individuals move from a feeling of protofusion to fusion. Also, an intercultural comparison of this particular period would be particularly interesting to understand the origin and causes of fusion.

Another issue that should be further explored is the causes of identity fusion. We know so far that sharing intense experiences with other ingroup members fosters fusion with the

group, particularly if these experiences are negative, traumatic, or dysphoric. Recent research has showed that positive experiences were also associated with identity fusion and pro-group actions (Kavanagh et al., 2018). On the other hand, taking together research showing that encouraging fused persons to focus on the shared core values of ingroup members increase their endorsement of making extreme sacrifices for such group (Swann, Buhrmester et al., 2014), future research should explore the relative influence of shared experiences, positive and negative, and shared core values on identity fusion and its correlates.

An extremely promising new line of research is to explore the underlying mechanisms involved when individuals are fused with an entity different from the group, as another individual, a value, an animal, a brand, etc. Researchers should not assume that the antecedents, consequents, mediators, and moderators are the same that those that have been found in the literature when researchers have focused on being fused with a group. For instance, it is evident that people cannot develop relational ties with things that have not a mind (e.g., a brand, a religion). Consequently, the principle of relational ties, which is essential in fusion theory, does not apply to these entities. Future research should clarify whether there could be differences between fusion with groups and fusion with other entities regarding all these factors.

Considering all the misconceptions about identity fusion, the most persistent over time may be the assumption that fusion is merely super-identification. To disentangle this confusion, defenders of fusion have strived to demonstrate what identity fusion can predict and identification cannot, or to prove that fusion is a better or stronger predictor than identification of some particular phenomena. However, here we endeavour advocates of identity fusion and group identification to follow an alternative strategy, as it is to show those phenomena that can be predicted by identification and not by fusion. Findings in this line would be extremely helpful to distinguish the nature of these two mechanisms.

In the same line, something that would help reinforce identity fusion theory would be to distinguish fusion from other factors that could also predict extreme pro-group behaviour. As we have recognized previously, the originators of the theory never assumed that identity fusion is the single or the most powerful predictor of self-sacrificial behaviour. Then, other studies should try to disentangle the role played by fusion and other factors (e.g., values, individual differences) in the determination of extreme forms of pro-group behaviour. As we mentioned earlier, individuals who are not fused might perform self-sacrificial behaviour, and fusion can interact with other factors (e.g., threat, values) to amplify pro-group actions. In fact, there are already several theoretical proposals that integrate fusion with other determinants of extreme behaviour such as the devoted actor model (Gómez et al., 2017; Sheikh et al., 2016; Vázquez et al., 2020) and the 3N model (Belanger et al., 2019; Kruglanski, Bélanger et al., 2019; Webber & Kruglanski, 2016). These proposals are still in an initial phase of validation, but in the near future they will surely improve our understanding of extremism and radicalization.

Any of these questions constitutes a good starting point for future research. With the present manuscript, we want to help scientists interested in identity fusion to reach a better understanding of the nature of the construct, the most common misunderstandings that gravitate around it, and some of the most pressing questions that still remain unanswered. It is our hope that some of them take the lead of exploring these questions and contribute to the theoretical and practical development of identity fusion theory.

Conclusions

Unravelling what leads some people to sacrifice for others, or for a cause, has undoubtedly been the focus of interest of social scientists. A decade ago, a fresh theory, identity fusion, proposed that a visceral feeling of oneness with the group could help disentangle why some individuals are willing to fight, kill, or die for such group. Despite criticisms,

misunderstandings, and untested assumptions, the theory has generated enormous international, multidisciplinary interest not only for scientists, but also for governments. Here we have considered these appreciations, confusions, and unverified suppositions as a challenge rather than a threat. As a consequence, our responses have tried to generate future lines of research that could reinforce the theory and make it even more ambitious. Only five years after the appearance of Fredman et al.'s (2015) work in this same outlet, identity fusion is nowadays a fundamental mechanism for anyone interested in the investigation of violent radicalization and de-radicalization[2].

[2] We intended to be as exhaustive as possible in our review of the main advances and discoveries about identity fusion. Although all contributions are valuable, word limitations forced us to exclude some papers from it. We would like to apologize for these omissions.

CHAPTER 3

Admiration for Islamist Groups Encourages Self-sacrifice through Identity Fusion

Reference

Gómez, Á., Bélanger, J., **Chinchilla, J.,** Vázquez, A., Schumpe, B. M., Nisa, C. F., & Chiclana, S. (2021). Admiration for Islamist groups encourages self-sacrifice through identity fusion. *Humanities & Social Sciences Communications, 8*(54). https://doi.org/10.1057/s41599-021-00734-9.

Abstract

The psychological mechanisms that lead terrorists to make costly sacrifices for their ideological convictions are of great theoretical and practical importance. We investigate two key components of this process: (1) the feeling of admiration toward ingroup members making costly self-sacrifices for their ideological group, and (2) identity fusion with religion. Data collected in 27 Spanish prisons reveal that Jihadists' admiration toward members of radical Islamist groups amplifies their willingness to engage in costly sacrifices for religion in prison. This effect is produced because admiration toward radical Islamist groups has a binding effect, increasing identity fusion with religion. Five additional experiments provide causal and behavioural evidence for this model. By showing that admiration for ingroup members increases identity fusion, which in turn makes individuals prone to engage in costly pro-group behaviours, we provide insights into the emotional machineries of radicalization and open new avenues for prevention strategies to strengthen public safety.

Keywords: Admiration, identity fusion, Jihadism, self-sacrifice, religion

Introduction

Admiration is a social emotion that is evoked by observing laudable behaviours – such as generosity, loyalty, fidelity, and commitment (Immordino-Yang et al., 2009; Pohling and Diesner, 2016; Schindler et al., 2013) – in others. In group dynamics, admiration can lead members who marvel at the conviction of their peers to become inspired to engage in similar, if not costlier, behaviours to support their group. For groups engaged in violence, this contagious effect can have dramatic consequences. Two contrasting examples illustrate this point. On March 15, 2019, a person previously unknown to authorities attacked two mosques in Christchurch, New Zealand, killing 51 people and injuring 49. Australian Prime Minister Scott Morrison described the perpetrator as an "extremist, right-wing, violent terrorist" with an admiration for groups supporting violence against Muslims.

Less than three years earlier, up to 30,000 foreign fighters from 100 countries joined the Islamic State (ISIS) following its proclamation of an Islamic caliphate forged through mass executions, genocide, and suicide bombings (UN, 2019). Among them, thousands of indoctrinated child soldiers ("cubs of the caliphate") are spurred into action through the lure of status, purpose, and admiration (Nyamutata, 2020).

But while these examples illustrate the range of political behaviours admiration can beget, they do not explain why admiration for an individual or a group produces greater willingness to engage in costly sacrifices. What is it about admiration that makes individuals willing to risk life and limb on the altar of an ideology? Answering this question and understanding the processes that drive active involvement in ideological groups has far-reaching social and geopolitical ramifications, and may help to inform policy decisions to prevent radicalization.

In the present research, we propose that identity fusion – an individual's visceral commitment to a group or ideology – is what binds admiration and self-sacrifice together. In

particular, we suggest that admiration viscerally connects people to their group or their convictions, which in turn leaves them prone to making costly sacrifices for their deep-seated beliefs. To examine this emotion-driven response, we report on fieldwork conducted with Jihadists (ISIS and al-Qaeda supporters) incarcerated in Spain, with a control group of Muslims imprisoned for crimes unrelated to terrorism (Study 1) and combine this field study with five online experimental studies offering causal relation and refinement to our model tested in prisons.

Literature Review

According to contemporary emotion theories, admiration is an other-focused emotion (Algoe & Haidt, 2009; Onu et al., 2016) that hinges on having "regard for someone or something considered praiseworthy or excellent" (Oxford English Dictionary, 2020). Admiration is often described as a transfiguring experience bestowing a kind of enlightenment leading to important personal changes (Pohling & Diessner, 2016). It is important to highlight that admiration is conceptually different from other social emotions that are positively valenced such as gratitude, awe, and adoration, which are associated with different elicitors and action tendencies. Specifically, whereas admiration is evoked by witnessing manifestations of virtue or excellence, gratitude occurs when someone is beneficiary of another's moral excellence (Onu et al., 2016). Awe also originates from witnessing virtue or excellence, but the event is perceived as so extraordinary and overwhelming that it is beyond one's comprehension (Keltner & Haidt, 2003). Furthermore, whereas admiration motivates emulation and self-improvement, gratitude motivates paying back the benefactor, and awe induces a state of contemplation and submission (Onu et al., 2016). And while admiration motivates the internalization and emulation of ideals embodied by an outstanding role model, "adoration stimulates adherence to the teachings and expectations of a meaning maker and benefactor perceived as superhuman or sacred" (Schindler et al., 2013).

Although relatively little empirical work has been conducted on admiration, some scholars use the label *moral elevation* for admiration elicited by acts of virtue and restrict use of the term admiration to emotional responses to non-moral excellence (Algoe & Haidt, 2009; Haidt, 2000). While these two concepts differ in some respects, moral elevation and admiration are both associated with outcomes relevant to the present research including self-improvement, the creation of social bonds, and the emulation of ideals embodied by outstanding exemplars (Algoe & Haidt, 2009). Therefore, given that moral elevation and admiration energise people to approach successful others and learn from them (Henrich & Gil-White, 2001; Smith, 2000), we postulate that both should produce identity fusion. Moreover, to contribute to this incipient literature, we also test whether admiration evoked by moral virtue produces greater identity fusion than admiration evoked by personal skills.

Theoretical explanations of the evolutionary origins and functions of admiration are rooted in the gregarious nature of human beings and their need to belong (Baumeister & Leary, 1995). The social functional approach to the study of emotions, proposes that admiration toward ingroup members creates greater group cohesion and devotion to its members (Stellar et al., 2017). According to multilevel selection theory, admiration motivates individuals to transcend their narrow personal interests and to make sacrifices to help the group achieve its goals (Pohling & Diessner, 2016). In fact, there is evidence suggesting that individuals' readiness to self-sacrifice for a cause is positively related to admiring exemplars that engage in costly behaviours (Bélanger et al., 2014). A large body of evidence also indicates that admiration promotes prosocial behaviours (Schnall et al., 2010; Vyver & Abrams, 2015) and behaviours that increase the welfare of others (Penner et al., 2005), such as charitable donations and volunteering (Cox, 2010; Schnall et al., 2010; Vianello et al., 2010; Vyver & Abrams, 2015). Likewise, there is a positive correlation between employees' admiration for their organizational leaders and self-reported feelings of altruism, commitment, courtesy, and

compliance at work (Vianello et al., 2010). A specific example of this view shows a relationship between Hong Kong residents' feelings of admiration toward pro-democracy protesters persecuted by the Chinese government, and their willingness to take political actions to challenge the government's anti-democratic policies (Sweetman et al., 2013).

However, despite indications that admiration is associated with self-sacrifices for the sake of a group, the scientific literature has failed to show any evidence, experimental or otherwise, explaining this relationship. We postulate that the core explanation of this phenomenon is identity fusion with the group.

Identity fusion is a visceral feeling of oneness with a group, marked by strong allegiance to each group member ("relational ties") and to the group as a whole including their goals and values ("collective ties"; Gómez et al., 2019; Swann et al., 2012). Together, these attachments trigger feelings of agency that motivate pro-group actions. For strongly fused individuals, the borders between personal identity (the part of our identity that derives from the characteristics that make us unique) and social identity (the part that stems from our membership in groups) become porous (Swann et al., 2012).

Furthermore, when fused individuals engage in pro-group behaviours, they experience a profound sense of personal agency that they put at the service of the group's wellbeing. Research has also shown that activating personal or social identities produces similar responses – namely, a greater willingness to fight and die for the group (Gómez, Brooks et al., 2011; Swann et al., 2009). Moreover, fused individuals are strongly attracted to other group members; they value them as if they were brothers and sisters. Once individuals are fused with a group, they tend to remain fused (Gómez et al., 2020; Swann et al., 2012).

Most recent theories on radicalization, such as the devoted actor model and the 3N model of radicalization (Bélanger et al., 2018; 2019; Gómez et al., 2017; Kruglanski, Bélanger et al., 2019; Kruglanski, Webber et al., 2019; Sheikh et al., 2016; Vázquez et al., 2020),

recognise identity fusion as an important psychological factor connecting individuals to radical networks, and in predicting one's willingness to engage in violence on the group's behalf. The devoted actor model proposes that individuals fused with a group with which they share a sacred value are particularly willing to display costly sacrifices for the group and the value (Gómez et al., 2019). The 3N model of radicalization, on the other hand, posits that the need for personal significance (i.e., the desire to matter, to experience a meaningful existence) is related to a "collectivistic shift" orienting individuals towards joining and strongly identifying with a group, which may lead to violent extremism if the ideological narrative of the group elevates violence as a means of attaining significance (Bélanger et al., 2018; 2019; Kruglanski, Bélanger et al., 2019; Kruglanski, Webber et al., 2019). In this process, members of the group validate the narrative by respecting and venerating those who engage in the requisite violence against the enemies of the group.

Identity fusion has also been linked to behaviours such as the refusal to leave the group (Gómez, Morales et al., 2011); denial of group wrong-doing (Besta et al., 2014); diminished quality of life after one's group is defeated (Buhrmester et al., 2012); the willingness to participate in extreme forms of protest on the group's behalf (Kunst et al., 2018); maximizing ingroup advantage over an outgroup, even at one's personal expense (Buhrmester, Burnham et al., 2018); protecting group reputation (Ashokkumar et al., 2019); sacrificing personal relationships (Swann et al., 2015); relative intergroup formidability (Buhrmester et al., 2014); donating to charity (Buhrmester et al., 2014; Swann et al., 2010); writing supportive notes to victims of terrorist attacks (Buhrmester et al., 2014); behaving aggressively toward outgroup members in videogames (Vázquez et al., 2020); and the desire to retaliate against outgroup members (Fredman et al., 2017). Recently, research has found that identity fusion also extends to other targets such as ideological convictions, other individuals, animals, objects, or even activities (Gómez et al., 2020).

Although the relationship between identity fusion and costly pro-social behaviours has been widely documented, research on the antecedents of identity fusion is scarcer.

Two main factors have been proposed as potential causes of identify fusion: 1) shared biology and 2) shared experiences. There is ample evidence to support both approaches.

The first cause – shared biology – is backed by data showing that monozygotic twins are more fused with their sibling than dizygotic twins (Vázquez et al., 2017), and are also more willing to die for their sibling (Tornero et al., 2017). Several investigations have demonstrated that the family is the group for whom individuals are most willing to die – a robust finding shown in 11 countries across five continents (Swann, Buhrmester et al., 2014), including an investigation with Libyan revolutionaries during the 2011 conflict in Libya (Whitehouse et al., 2014), a study conducted with frontline fighters against ISIS (Gómez et al., 2017), and a study conducted with young men in Mosul, Iraq (Atran, 2019). Experimental research has also shown that priming the existence of shared biological attributes between members of a group increases the perception of family-like ties, which, in turn, increases people's willingness to engage in costly sacrifices for the sake of the group (Swann, Buhrmester et al., 2014).

The second possible cause – shared experiences – is also well documented. Several investigations have offered empirical evidence that sharing intense experiences with ingroup members fosters fusion with the group (Kapitány et al., 2018; Kavanagh et al., 2018; Misch et al., 2018), when such experiences are positive (Kavanagh et al., 2018), but particularly if these experiences are negative, traumatic, or dysphoric (Jong et al., 2015; Newson et al., 2017; Segal et al., 2018; Whitehouse et al., 2017). Similarly, participating in collective gatherings, like folkloric marches and religious celebrations, could increase fusion (Páez et al., 2015; Zumeta et al., 2016), while engaging in ritual practices or recalling episodic memories of pilgrims' routes contribute to maintaining identity fusion (Lobato & Sainz, 2019).

Beyond these findings, however, research has been silent on the role of social emotions in encouraging identity fusion. In the present research we extend and integrate previous work on emotions, identity fusion, and violent extremism by proposing that admiration increases costly sacrifices for an ideology through greater identity fusion.

To test the validity and real-world relevance of our model, we first report data from a quantitative field study conducted with imprisoned Jihadists in Spain that compares their responses to a control group of criminals (also Muslims) imprisoned for crimes unrelated to terrorism. Following this unique field study, we present five experimental studies to ensure that our claims are causal and robust (see Table 2 for overview of studies and research questions addressed).

The study with prisoners and the online studies were approved by the Institutional Review Board of the authors' university and conducted with the permission of the prison's authority. All studies were conducted in accordance with the rules and regulations of the American Psychological Association. Informed consent was obtained from all participants before they took part in the studies. Participants anonymity was assured along with explicit assurance that interviews or experiments involving verbal answers to questionnaires could be terminated at will. Participants agreed on a voluntary basis to participate in interviews and experiments involving answers to questionnaires. In accordance with standard IRB protocols, they were reminded that they could abandon the investigation at any time if they felt uncomfortable. Participants were thanked and debriefed at the end of the interview and questionnaire. They received the contact information of the PI in case they were interested in receiving further information.

Table 2. Overview of studies and specific research questions addressed

	Main Goal	**Experimental Conditions**	**Locus of Admiration**	**Locus of Fusion**	**Outcome measure**
Study 1	Initial test of paradigm	—	Radical Islamist Groups	Religion	Sacrifices for Religion in Prison
Study 2	Testing model with experimental manipulation of admiration	Admiration vs. No Admiration	Fellow Citizen	Country (i.e., Spain)	Willingness to Fight and Die for the Country
Study 3	Testing if the model is moderated by the number of ingroup members being admired	2 (Admiration vs. No Admiration) X 2 (Individual vs. Group)	A Fellow Citizen vs. Group of Citizens	Country (i.e., Spain)	Willingness to Fight and Die for the Country
Study 4	Testing if the model is moderated by the locus of admiration (i.e., ingroup vs. outgroup)	2 (Admiration vs. No Admiration) X 2 (Ingroup vs. Outgroup)	Fellow Citizen vs. Citizen from other Country	Country (i.e., Spain)	Willingness to Fight and Die for the Country
Study 5	Testing whether the model is moderated by the reason for experiencing admiration	2 (Admiration vs. No Admiration) X 2 (Personal Qualities vs. Sacrifices for the Group)	Fellow Citizen	Country (i.e., Spain)	Willingness to Fight and Die for the Country
Study 6	Testing whether admiration 1) produces change in behaviour 2) one month after experiencing it.	Admiration vs. No Admiration	Fellow Citizen	Country (i.e., Spain)	Prosocial/Aggressive Behaviour toward Ingroup Members

Study 1: Admiration towards Radical Islamist Groups Increases Identity Fusion with Religion and Willingness to Engage in Costly Sacrifices among Imprisoned Jihadists

We tested whether imprisoned Jihadists (vs. non-terrorist criminals) would report greater admiration towards radical Islamist groups, which in turn would be associated with greater identity fusion with religion, and greater willingness to engage in costly sacrifices for religion in prison. The study was conducted in Spain. The Spanish Penal code (articles 571 to 580) defines terrorism as crimes carried out for any of the following purposes: (1) Subvert the constitutional order, or seriously suppress or destabilise the functioning of the political institutions or the economic or social structures of the State, or compel the public powers to carry out an act or to refrain from doing so; (2) seriously alter public peace; (3) seriously destabilise the cooperation of an international organization; and (4) provoke a state of terror in the population or in part of it. Terrorist acts include crimes against the Crown, life, physical or moral integrity, liberty, heritage, public health, and natural resources or the environment. Possession, traffic and deposit of weapons, ammunition, or explosives, and the seizure of aircrafts, ships, or other means of transport are also included under the umbrella of terrorism. Finally, computer intrusion and computer damage are typified as terrorism as well when they serve some of the above stated goals.

Method

Participants

We selected two sub-samples of prisoners who participated in a broader research and that responded to the questions prepared for the goal of the present investigation. One of the subsamples included returning ISIS foreign fighters, participants in the 2004 and 2017 attacks in Madrid and Barcelona, respectively, and others convicted of plotting terrorist attacks for

Islamist causes. The other subsample included Muslims involved mainly in petty criminal activities (drug trafficking, robbery, etc.). Prisoners in this research were all men and Muslims. Thirty-six jihadists (M_{age} = 32.81, SD_{age} = 8.80) and 35 non terrorist criminals (M_{age} = 35.77, SD_{age} = 8.70) located in 27 Spanish prisons took part in the study.

Procedure

Approximately one week before the interviews, prisoners were asked if they would like to participate voluntarily in a research about life conditions in prison, their feelings and emotions related to different groups they belong to, and their values. The day of the interview, participants were individually interviewed by trained members of our research team in a private room within the prison complex which was habilitated for the purpose of the study by the prison staff.

At the beginning of the interview, participants were informed that all the responses would be anonymous, that they were free to not respond to any of the questions, and that no benefit or changes to their sentence or conditions would follow no matter if they chose to participate or decline to take part in this study. After taking part in a semi-structured interview and answering some questions unrelated to the purposes of the study, participants responded to a series of items measuring their admiration toward radical Islamist groups, their fusion with religion, and their willingness to engage in costly sacrifices for religion in prison. In all the following studies, responses to the measures of perceived cost of sacrifice and willingness to fight and die were scored on Likert-type scales ranging from 1 (*totally disagree*) to 7 (*totally agree*), whereas responses to the rest of the measures were scored on scales ranging from 0 (*totally disagree*) to 6 (*totally agree*).

Feelings of admiration were measured with a single item. We asked each participant to "indicate to what degree you admire the members of Islamist activist/radical groups considering that 0 means "I do not admire them at all" and 6 means "I admire them a lot."

Fusion with religion was measured using the Dynamic Identity Fusion Index (DIFI; Jiménez et al., 2016). This measure displays a figure made of two circles of different sizes on the computer screen. A small circle on the left side of the computer screen represents "the self." A bigger circle on the right side of the screen represents "religion". Participants were asked to move the small circle until it reached the position that best represented their relationship with religion. Higher degrees of overlap of both circles, from 0 to 100, indicate higher levels of fusion. *Costly sacrifices for religion* were measured with a five-item scale developed specifically for the study supervised by professional psychologists that work in prison (for example, "If it was necessary, I would be willing to move to a prison further away from my family to defend my religion", $\alpha = .96$), (see Supplementary Information).

Results and Discussion

Table 3 presents the correlations between the variables included in the different studies.

Table 3. Bivariate correlations for Studies 1 to 6

		Admiration	Sacrifices	
Study 1	Fusion	.37**	-	
	Sacrifices	.48**	.54*	
		Cost of sacrifice	Fusion	
Study 2	Fusion	.46**	-	
	Fight-die	.39**	.56***	
Study 3	Fusion	.31**	-	
	Fight-die	.29**	.60**	
Study 4	Fusion	.27**	-	
	Fight-die	.35**	.53**	
Study 5	Fusion	.33**		
	Fight-die	.31**	.58**	
		Cost of sacrifice	Fusion	Help
Study 6	Fusion	.36**	-	
	Help	.26**	.22**	-
	Aggression	-.23**	-.20**	-.77**

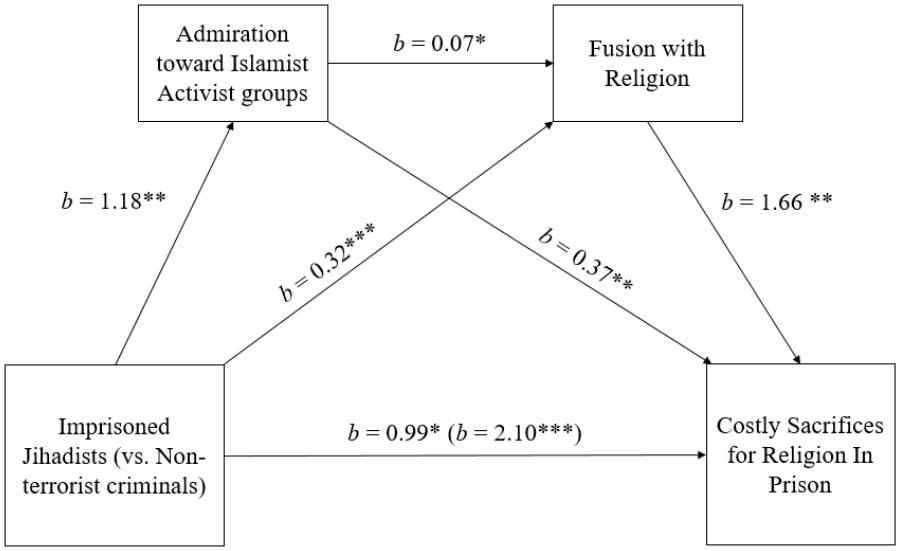
Note. ** $p < .01$; *** $p < .001$

We performed a multivariate analysis of variance (MANOVA) on admiration, identity fusion, and costly sacrifices. In line with our predictions, results indicated that Jihadists (vs. non-terrorist criminals) reported more admiration toward radical Islamist groups, $F(1, 68) = 9.27$, $p = .003$, $\eta^2_p = .12$, Ms = 0.91 vs. 0.11, SDs = 1.90 vs. 0.53, more fusion with religion, $F(1, 68) = 17.32$, $p < .001$, $\eta^2_p = .20$, Ms = 63.87 vs. 26.29, SDs = 40.94 vs. 38.93, and more willingness to engage in costly sacrifices for religion in prison, $F(1, 68) = 19.41$, $p < .001$, $\eta^2_p = .222$, Ms = 2.61 vs. 0.61, SDs = 2.43 vs. 1.27.

To test our hypothesis that feelings of admiration toward radical Islamist groups and fusion with religion mediate the effect of groups on willingness to perform costly sacrifices for religion in prison, we performed a mediation analysis using Hayes' (2018) PROCESS Macro, Model 6. We included the group (0 = non-terrorist *criminals*, 1 = *Jihadists*) as the predictor, admiration as the first mediator, identity fusion as the second mediator, and costly sacrifices as the outcome variable. This analysis provided support for the proposed mediation model: Jihadists (vs. non-terrorist criminals) reported greater willingness to engage in costly sacrifices for religion in prison because of greater admiration toward radical Islamist groups and fusion with religion (see Figure 1).

While Study 1 offers quasi-experimental evidence for our model, Studies 2 to 6 were conducted to present causal evidence through a series of experimental and longitudinal studies. In these studies, we also examined in greater detail why admiration produces greater identity fusion and self-sacrifices. For the next studies, power analyses were conducted with 5,000 Monte Carlo simulations. Assuming small-to-medium effect sizes and power set at .80, a sample of 135 participants were recommended in Studies 2 and 6. A sample size of 215 people were recommended in Studies 3, 4, and 5.

Figure 1. Lineal mediation model in Study 1 ($N = 171$)



Jihadists expressed more willingness to engage in costly sacrifices for religion in prison via 1) admiration toward radical Islamist groups alone, $b = 0.44$, s.e. = 0.25, 95% CI [0.02, 0.97], 2) fusion with religion alone, $b = 0.54$, s.e. = 0.29, 95% CI [0.13, 1.22], and 3) admiration and fusion serially, $b = 0.13$, s.e. = 0.07, 95% CI [0.02, 0.30], (Study 1).
Note. * $p < .05$; ** $p < .01$; *** $p < .001$

Study 2: Feelings of Admiration Increase Perceived Intensity of Self-sacrifice, Identity Fusion, and Willingness to Fight and Die

In line with our theoretical framework that admiration is an emotion elicited by virtue (Cox, 2010; Snall et al., 2010; Vianello et al., 2010; Vyver & Abrams, 2015), we hypothesised that admiring someone increases the saliency of the costly sacrifices that person made for the group, such that the more intense the perceived sacrifices, the greater the identity fusion with the group, and the willingness to engage in costly behaviours.

Method

Participants

Participants for this study and the rest of the online studies were recruited using a snowball procedure wherein Psychology students invited their acquaintances to volunteer.

One hundred and fifty-two participants volunteered for this study (M_{age} = 39.62, SD_{age} = 13.89, 65.1% women).

Procedure

The study used an experimental design. After obtaining participants' informed consent, participants were randomly assigned to one of two experimental conditions. In the *admiration condition* (n = 76) participants were asked to think about a person from their country they admire and write about who this person is and the reasons for admiring him/her. In the *typical condition* (n = 76), participants were asked to think about a typical citizen of their country and to describe that person. After that, participants completed our outcome variables and the manipulation check.

As a manipulation check, *feelings of admiration* were measured with a four-item scale developed for the study (for example, "I admire this person's actions", α = .91) in all studies. The results confirmed that our manipulations were successful (see Supplementary Information).

Cost of personal sacrifice was measured with a five-item scale developed for the study (for example, "This person incurs a great loss to further the interests of my country", α = .78).

Fusion with the country was measured with the seven-item verbal measure of fusion developed by Gómez, Brooks et al. (2011; for example, "I am one with my country", α = .86).

Willingness to fight and die for the country was measured with a seven-item scale (Swann et al., 2019; for example, "Hurting other people is acceptable if it means protecting my country", α = .83).

Lastly, participants' level of fusion with groups unrelated to the manipulation (their family, France, Portugal, and Italy) was measured with the DIFI adapted to these groups. In this and the following studies, the effects of our manipulations on fusion with groups different

from Spain were not significant, showing that the strategies that we used to induce admiration did not increase identity fusion indiscriminately (see Supplementary Information).

Results and Discussion

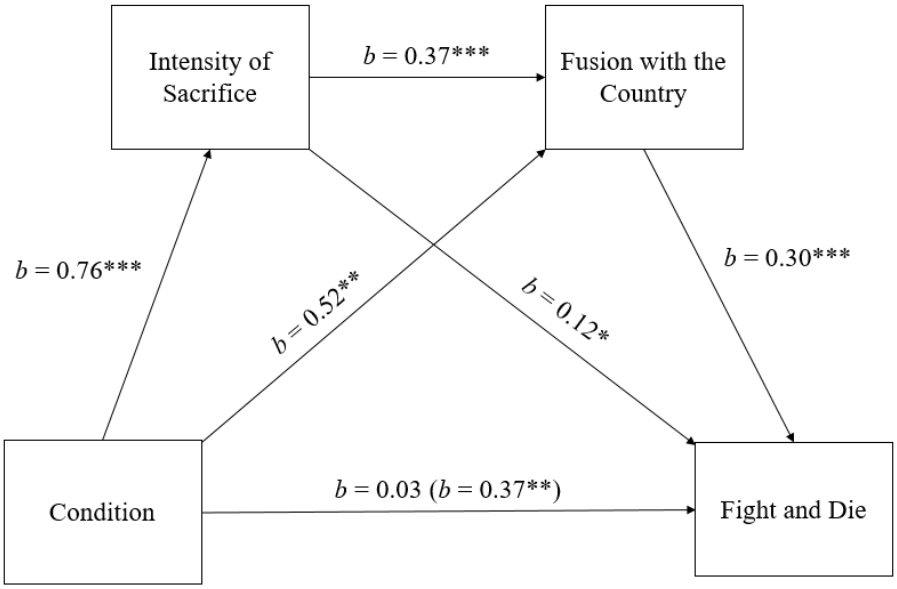
A multivariate analysis of variance (MANOVA) showed that participants in the admiration (vs. typical) condition perceived the ingroup member's sacrifices as more intense, $F(1, 150) = 16.31$, $p < .001$, $\eta^2_p = .10$, felt more fused with their country, $F(1, 150) = 18.77$, $p < .001$, $\eta^2_p = .11$, and were more willing to fight and die for it, $F(1, 150) = 7.78$, $p = .006$, $\eta^2_p = .05$ (see Table 4 to obtain more information about the means and standard deviations per condition of Studies 2-6).

Table 4. Means and standard deviations per condition for Studies 2 to 6

Study	Experimental Condition	Cost of Sacrifice		Fusion		Fight and Die		Help		Aggression	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
2	Admiration	3.74	1.25	2.59	1.24	2.11	0.94				
	Typical	2.98	1.08	1.78	1.05	1.74	0.66				
3	Admiration Individual	3.69	1.40	2.40	1.22	2.22	1.09				
	Admiration Group	4.15	1.26	2.39	1,34	2.09	1.15				
	No Admiration Individual	2.12	0.72	1.89	1.11	1.85	0.72				
	No Admiration Group	2.43	0.68	1.89	1.26	1.71	0.76				
4	Admiration Ingroup	3.91	1.27	2.50	1.34	2.12	1.28				
	Admiration Outgroup	3.50	1.34	2.01	1.15	1.73	0.66				
	No Admiration Ingroup	2.20	0.91	1.63	0.96	1.50	0.41				
	No Admiration Outgroup	2.50	1.07	1.73	1.11	1.67	0.57				
5	Admiration Sacrifices	4.23	1.11	2.93	1.14	2.28	0.82				
	Admiration Qualities	3.26	1.36	2.03	1.25	1.88	0.91				
	No Admiration Sacrifices	2.36	0.97	1.78	1.29	1.82	0.80				
	No Admiration Qualities	2.28	0.96	2.26	1.32	1.94	0.81				
6	Admiration	3.90	1.19	2.82	1.11			6.18	2.87	0.70	1.17
	No Admiration	2.55	0.95	1.92	0.99			4.13	2.71	1.51	1.66

To test our hypothesis that cost of sacrifice and identity fusion serially mediate the effect of the experimental condition on willingness to fight and die for the country, we used PROCESS macro, Model 6. Condition (0 = *typical*, 1 = *admiration*) was included as the predictor, cost of sacrifice as the first mediator, identity fusion as the second mediator, and willingness to fight and die as the outcome variable. Results indicated that admiration for an ingroup member produced greater willingness to fight and die for the country through its influence on the perceived intensity of sacrifices and identity fusion (see Figure 2). In the next study we explored whether admiration toward a group of people rather than specific individuals would produce similar effects.

Figure 2. Lineal mediation model in Study 2 ($N = 152$)



Admiration for an ingroup member increased willingness to fight and die via 1) intensity of sacrifices alone, $b = 0.10$, 95% CI [0.03, 0.20], 2) identity fusion alone, $b = 0.16$, 95% CI [0.05, 0.30], and 3) intensity of sacrifice and fusion with country serially, $b = 0.08$, 95% CI [0.03, 0.17]. (Study 2).
Note. * $p < .05$; ** $p < .01$; *** $p < .001$

Study 3: Feelings of Admiration toward an Individual Versus a Group

Study 3 was conducted to clarify whether admiration toward a group of fellow countrymen or admiration toward a single compatriot, increases willingness to fight and die for the country through perceived intensity of sacrifices and identity fusion.

Method

Participants

Two hundred and thirty-one participants (M_{age} = 39.76, SD_{age} = 13.69, 61.9% women) volunteered for this study.

Procedure

Study 3 featured an experimental 2 (no admiration vs. admiration) x 2 (individual vs. group) design. After obtaining participants' informed consent, participants were randomly assigned to one of four experimental conditions. In the *individual admiration condition* (n = 52), participants were asked to think about a person from their country they admire and write about him/her as in Study 2. In the *no admiration individual condition* (n = 59), participants were asked to think about a person of their country they do not admire and to write about who this person is and the reasons for their lack of admiration. Participants in the *group admiration condition* (n = 65), and the *group no admiration condition* (n = 55) received similar instructions, but they were asked to think about a group of people from their country instead of a person. Then, participants were asked to complete the same measures as in Study 2 adapted to the individual or group conditions: *admiration* toward the person or group they thought about (α = .96), *perceived cost of the personal sacrifice* (α = .66), their *fusion with the country* (α = .87), *willingness to fight and die for the country* (α = .84), and *fusion with other groups*.

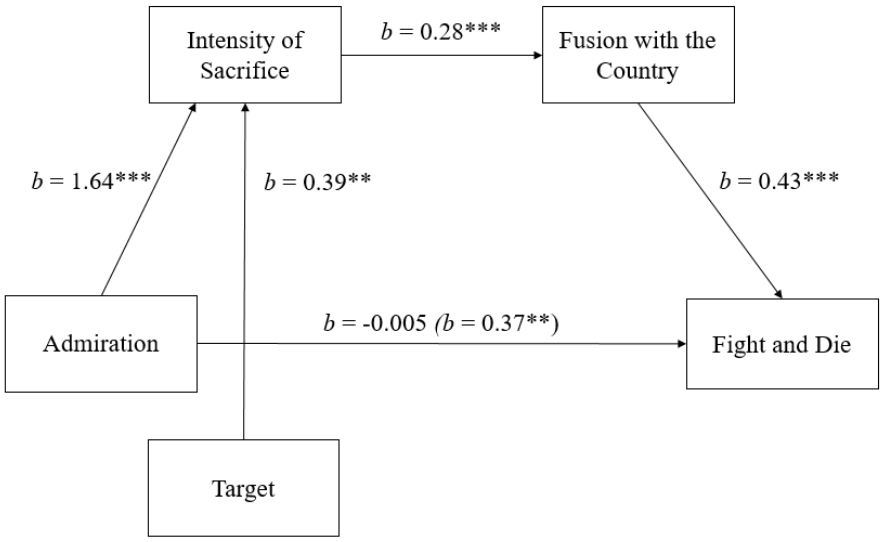
Results and Discussion

A 2 (no admiration vs. admiration) x 2 (individual vs. group) MANOVA on the outcome measures yielded significant main effects of admiration. In the admiration (vs. no admiration) conditions, participants perceived the ingroup member's sacrifices as more intense, $F(1, 227) = 135.98$, $p < .001$, $\eta^2_p = .37$, felt more fused with their country, $F(1, 227) = 9.38$, $p = .002$, $\eta^2_p = .04$, and were more willing to fight and die for it, $F(1, 227) = 8.48$, $p = .004$, $\eta^2_p = .04$. The analysis on cost of sacrifice also revealed a significant main effect of target of admiration indicating that participants in the group (vs. individual) conditions perceived the ingroup members' sacrifices as more intense, $F(1, 227) = 7.76$, $p = .006$, $\eta^2_p = .03$. No other effects were significant, *p*s = .305 (see the statistics for the main effects of all studies in Supplementary Information).

To test our mediational hypothesis that cost of sacrifice and identity fusion would serially mediate the effect of admiration on willingness to fight and die for the country, we performed a serial mediation analysis as in Study 2. As expected, admiration increased willingness to fight and die for the country through its influence on intensity of sacrifices and identity fusion.

Consistent with expectations from prior research on fusion (Gómez et al., 2011, 2020; Swann et al., 2009), the analyses showed the effects were not moderated by whether admiration was experienced toward a group or an individual (see Figure 3). However, the impact of admiration could vary depending on whether or not the admired person belongs to the ingroup.

Figure 3. Lineal mediation model in Study 3 ($N = 231$)



Admiration increased willingness to fight and die for the country via intensity of sacrifice and fusion serially, $b = 0.20$, 95% CI [0.09, 0.33]. However, admiration did not influence willingness to fight and die via intensity of sacrifices alone, $b = 0.16$, 95% CI [-0.004, 0.32], or identity fusion alone, $b = 0.01$, 95% CI [-0.14, 0.18].
Note. * $p < .05$; ** $p < .01$; *** $p < .001$

Study 4: Feelings of Admiration towards an Ingroup versus an Outgroup Member

Study 4 analysed whether the direct and indirect effects of admiration on willingness to fight and die for the group are moderated by the group membership of the admired person.

Method

Participants

Two hundred and fifty-three participants ($M_{age} = 38.11$, $SD_{age} = 13.41$, 68% women) volunteered for this study.

Procedure

Study 4 featured an experimental 2 (no admiration vs. admiration) x 2 (ingroup vs. outgroup) design. Participants were randomly assigned to one of four experimental conditions.

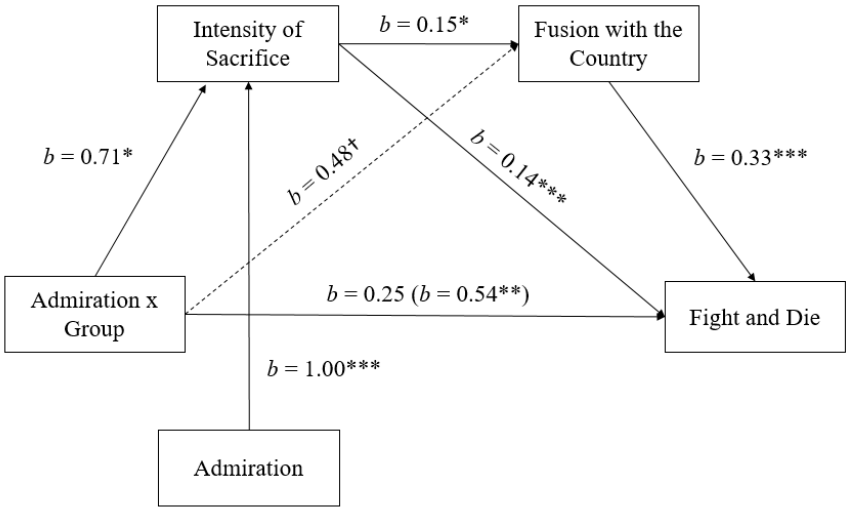
In the *ingroup admiration condition* (n = 68), participants were asked to think about a person of their country they admire. In the *ingroup no admiration condition* (n = 60), participants were asked to think about a person of their country they do not admire. In the *outgroup admiration condition* (n = 66), participants were asked to think about a person they admire from another country. In the *outgroup no admiration condition* (n = 59), participants were asked to think about a person of a different country they do not admire. In all cases, participants were asked to write about who that person is and the reasons for admiring (or not admiring) that person. Then, we measured participants' *admiration* toward the person they thought about (α = .94), the *perceived cost of the sacrifice* for the country attributed to him/her (α = .68), their *fusion with the country* (α = .84), their *willingness to fight and die for the country* (α = .82), and their *fusion with other groups*. These variables were assessed as in Studies 2-3.

Results and Discussion

A 2 (no admiration vs. admiration) x 2 (ingroup vs. outgroup) MANOVA revealed significant interactions on intensity of the sacrifices, $F(1, 249) = 5.80$, $p = .017$, $\eta^2_p = .02$, fusion $F(1, 249) = 4.17$, $p = .042$, $\eta^2_p = .02$, and willingness to fight and die for the country, $F(1, 249) = 6.99$, $p = .009$, $\eta^2_p = .03$. The effects of admiration were higher in the ingroup condition than in the outgroup condition: $F(1, 249) = 68.20$, $p < .001$, $\eta^2_p = .22$, vs. $F(1, 249) = 22.82$, $p < .001$, $\eta^2_p = .08$, for cost of the sacrifices; $F(1, 249) = 17.86$, $p < .001$, $\eta^2_p = .07$, vs. $F(1, 249) = 1.71$, $p = .193$, $\eta^2_p = .01$, for identity fusion; and, $F(1, 249) = 18.00$, $p < .001$, $\eta^2_p = .07$, vs. $F(1, 249) = 0.23$, $p = .633$, $\eta^2_p = .00$, for willingness to fight and die for the country. That is, participants in the admiration conditions perceived higher intensity of the sacrifices, felt more fused with their country and were more willing to fight and die for it than those in the no admiration conditions. However, those differences were weaker when participants thought about outgroup members than when they thought about ingroup members.

To test our hypothesis that cost of sacrifice and identity fusion mediate the interactive effect of admiration and group on willingness to fight and die for the country, we performed a moderated serial mediation analysis using PROCESS macro, Model 85. We included level of admiration (0 = *no admiration*, 1 = *admiration*) as the predictor, group (0 = *outgroup*, 1 = *ingroup*) as the moderator, cost of sacrifice and identity fusion (centred) as the first and second mediators respectively, and willingness to fight and die as the outcome variable. Results revealed that in the ingroup conditions, admiration influenced willingness to fight and die through its influence on intensity of sacrifice and fusion with country. In the outgroup conditions, the indirect effect was also significant, but weaker than in the ingroup conditions (see Figure 4).

Figure 4. Lineal mediation model in Study 4 (*N* =253)



In the ingroup conditions, level of admiration influenced willingness to fight and die via 1) intensity of sacrifice alone, b = 0.24, 95% CI [0.10, 0.40], 2) identity fusion alone, b = 0.20, 95% CI [0.06, 0.36], and 3) intensity of sacrifices and fusion with country serially, b = 0.09, 95% CI [0.01, 0.20]. In the outgroup conditions, level of admiration influenced willingness to fight and die via 1) intensity of sacrifice alone, b = 0.14, 95% CI [0.05, 0.25], and 2) intensity of sacrifices and fusion with country serially, b = 0.05, 95% CI [0.005, 0.12]. The indirect effect via identity fusion alone was not significant, b = 0.04, 95% CI [- 0.10, 0.17]. In the outgroup conditions, the indirect effect was also significant, but weaker than in the ingroup conditions, b = 0.10, 95% CI [0.02, 0.21].
Note. * $p < .05$; ** $p < .01$; *** $p < .001$

So far, we have shown that feeling of admiration for one or more ingroup members increases the willingness to engage in pro-group behaviours. In the next study, we examine whether admiring someone for a specific reason influence this motivational process.

Study 5: Feelings of Admiration Due to Self-sacrifice versus Personal Qualities

Study 5 examined whether admiration exerts different effects on pro-group behavioural tendencies depending on whether it is due to the personal qualities of the admired person or to his/her sacrifices for the country.

Method

Participants

Two hundred and forty-one participants (M_{age} = 36.58, SD_{age} = 13.52, 63.1% women) volunteered for this study.

Procedure

Study 5 featured an experimental 2 (no admiration vs. admiration) x 2 (personal qualities vs. personal sacrifices for the group) design. After obtaining participants' informed consent, participants were randomly assigned to one of four experimental conditions. In the *admiration* or *no admiration for personal sacrifices conditions* (*n*s = 42 and 65), participants were asked to think about a person of their country they admire or do not admire because of the sacrifices he/she makes for it. In the *admiration* or *not admiration for personal qualities conditions* (*ns* = 64 and 69), participants were asked to think about a person from their country they admire or do not admire due to their personal characteristics. In all cases, participants were asked to specify who this person is and the reasons for admiring or not that person. Then, participants were asked to complete measures of their level of *admiration* toward the person

they thought about (α = .95), the *perceived cost of the sacrifice* for the country attributed to him/her (α = .70), their *fusion with the country* (α = .87), their *willingness to fight and die for the country* (α = .78), and their *fusion with other groups* as in Studies 2-4.

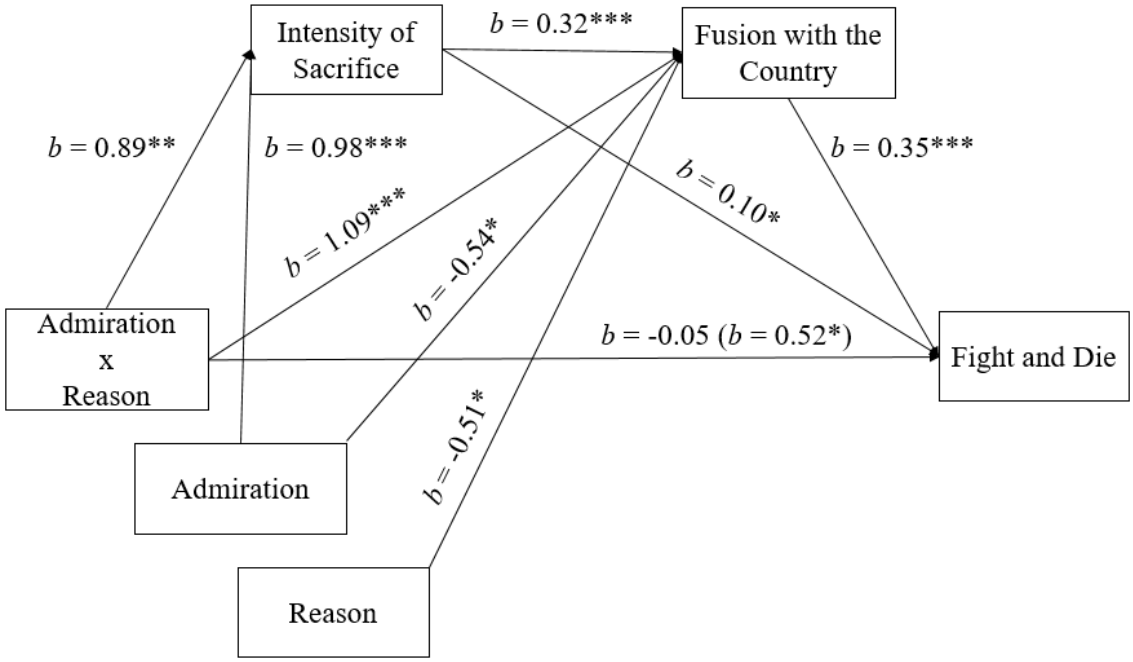
Results and Discussion

A 2 (no admiration vs. admiration) x 2 (personal qualities vs. personal sacrifices for the group) MANOVA revealed significant interactions on intensity of sacrifices, $F(1, 236) = 9.19$, $p = .003$, $\eta^2_p = .04$, identity fusion, $F(1, 236) = 17.07$, $p < .001$, $\eta^2_p = .07$, and willingness to fight and die for the country, $F(1, 236) = 5.54$, $p = .019$, $\eta^2_p = .02$. Specifically, the effects of admiration were greater in the personal sacrifices than in the personal qualities condition: $F(1, 236) = 71.88$, $p < .001$, $\eta^2_p = .23$, vs. $F(1, 236) = 25.71$, $p < .001$, $\eta^2_p = .10$, for cost of sacrifice; $F(1, 236) = 20.87$, $p < .001$, $\eta^2_p = .08$ vs. $F(1, 236) = 1.11$, $p = .292$, $\eta^2_p = .01$, for fusion with the country; and, $F(1, 236) = 7.95$, $p = .005$, $\eta^2_p = .03$ vs. $F(1, 236) = 0.13$, $p = .723$, $\eta^2_p = .00$, for willingness to fight and die for it. That is, participants in the admiration for personal sacrifices condition perceived higher intensity of the sacrifices, felt more fused with their country and were more willing to fight and die for it than those in the no admiration for personal sacrifices condition. However, those differences were weaker or non-significant when participants thought about personal qualities.

To test our hypothesis that cost of sacrifice and identity fusion would mediate the interactive effect of admiration and reason on willingness to fight and die for the country, we performed a moderated serial mediation analysis as in Study 4, but considering reasons of admiration (0 = *personal qualities*, 1 = *sacrifices*) as the moderator. Results revealed that in both the personal qualities and personal sacrifices conditions, admiration influenced willingness to fight and die through its influence on intensity of sacrifices and fusion with country. However, the effect in the personal qualities condition was smaller than in the personal

sacrifices condition (see Figure 5).

Figure 5. Lineal mediation model in Study 4 ($N = 241$)



In the personal sacrifice conditions, admiration influenced willingness to fight and die via 1) intensity of sacrifice alone, $b = 0.19$, 95% CI [0.02, 0.37], 2) identity fusion alone, $b = 0.19$, 95% CI [0.01, 0.38], and 3) intensity of sacrifices and identity fusion serially, $b = 0.21$, 95% CI [0.11, 0.33]. In the personal qualities conditions, admiration influenced willingness to fight and die via 1) intensity of sacrifice alone, $b = 0.10$, 95% CI [0.01, 0.21], 2) identity fusion alone, $b = -0.19$, 95% CI [- 0.35, - 0.04], and 3) intensity of sacrifices and identity fusion serially, $b = 0.11$, 95% CI [0.05, 0.19]. The effect in the quality's condition was smaller than in the personal sacrifice condition, $b = 0.09$, 95% CI [0.01, 0.21].
Note. * $p < .05$; ** $p < .01$; *** $p < .001$

Studies 1-5 examined the behavioural tendencies associated to admiration rather than actual behaviour. We solved this limitation in a final study conducted in two waves that additionally tested the temporal stability of the effects of admiration.

Study 6: Long-lasting Effects of Admiration on Real Behaviour

The last study extends previous results in two ways: it provides evidence of the effects of admiration over time and focuses on actual pro-group behaviour.

Method

Participants

Seven hundred and seventy-one participants (M_{age} = 36.26, SD_{age} = 12.91, 56.7% women) took part in the first wave and 376 participants (M_{age} = 37.29, SD_{age} = 13.59, 52.9% women) also completed the second wave.

Procedure

This study consisted of a longitudinal experimental design. In the first wave, participants were randomly assigned to the *admiration* (n = 391) or the no *admiration condition* (n = 380) as in Study 5 (personal sacrifices conditions). Then, we measured their feelings of *admiration* (α = .95), the *perceived cost of the sacrifice* for the country attributed to the person they thought about (α = .70), and their level of *fusion with their country* (α = .83) with the same scales used in Studies 2-5. We told participants that they would be contacted one month after wave 1 to participate in the second wave.

In the second wave, we used an adaptation of the tangram help/hurt task to measure participants' actual behaviour toward an ingroup member (Saleem et al., 2015). Participants were presented with a total of 30 different tangram puzzles that could be classified depending on their level of difficulty: 10 were easy, another 10 were difficult, and the remaining 10 were of medium difficulty. Participants were asked to assign 11 of these tangrams to another (fictitious) participant from their country, so that he/she could solve them in less than 10 minutes. We told participants that the other person could win a €25 voucher if he/she completed 10 tangrams in less than 10 minutes, otherwise the person wouldn't receive anything. Following the indications of Saleem et al. (2015), we calculated a measure of *helping behaviour*, operationalised as the number of easy tangrams assigned by participants greater than one, and a measure of *aggressive behaviour*, operationalised as the number of difficult tangrams assigned greater than one. Attrition analyses showed that participants who took part

in both waves were significantly less fused than participants who only took part in the first wave (see Supplementary Information).

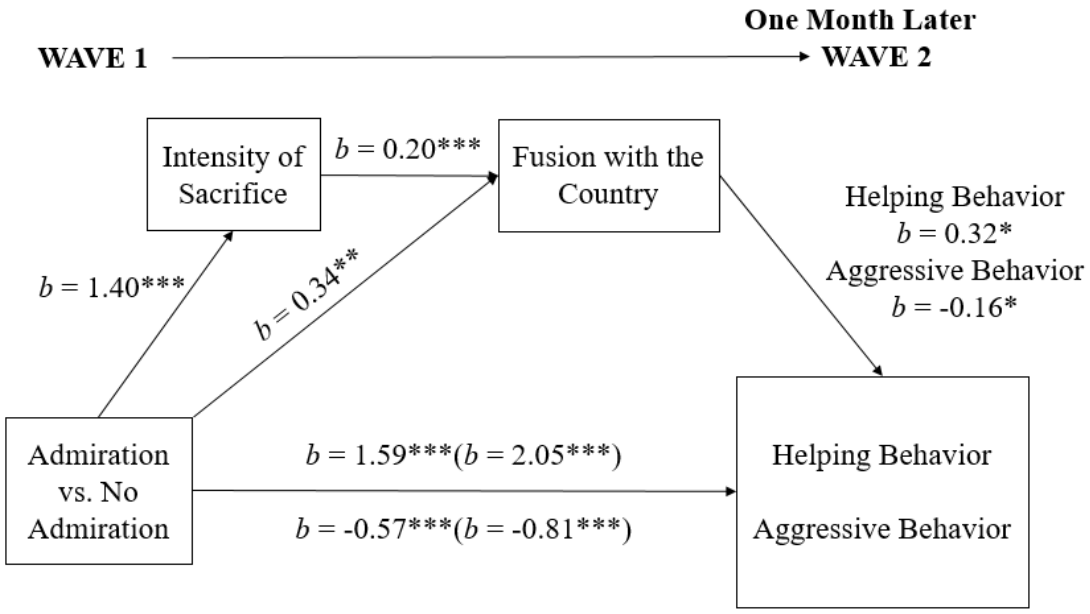
Results and Discussion

In the first wave, a MANOVA indicated that participants in the admiration (vs. no admiration) condition perceived the ingroup member's sacrifices as more intense, $F(1,769) = 303.76$, $p < .001$, $\eta^2_p = .28$, and were more fused with the group, $F(1,769) = 141.08$, $p < .001$, $\eta^2_p = .16$.

In the second wave, a mixed analysis of variance (ANOVA) with condition as a between-subject factor (0 = *no admiration*, 1 = *admiration*) and behaviour (helping vs. aggression) as a within-subject factor showed a significant condition X behaviour interaction, $F(1,374) = 48.34$, $p < .001$, $\eta^2_p = .11$. Results showed that participants engaged in more helping than aggressive behaviour, but such difference was higher in the admiration condition, $F(1, 374) = 383.87$, $p < .001$, $\eta^2_p = .67$ (*Mean difference* = 5.48, *SD* = 0.28, $p < .001$), than in the no admiration condition, $F(1,374) = 76.13$, $p < .001$, $\eta^2_p = .29$ (*Mean difference* = 2.63, *SD* = 0.30, $p < .001$). Follow-up analyses revealed that the effect of condition on helping behaviour was bigger (*Mean difference* = 2.05, *SD* = .29, $p < .001$), $F(1,374) = 50.45$, $p < .001$, $\eta^2_p = .12$, than the effect of condition on aggressive behaviour (*Mean difference* = -0.81, *SD* = .15, $p < .001$), $F(1,374) = 29.75$, $p < .001$, $\eta^2_p = .07$.

To test our hypotheses that intensity of sacrifice and identity fusion in Wave 1 would mediate the effects of condition on helping and aggressive behaviour in Wave 2, we performed two mediation analyses using PROCESS macro, Model 6. For both – helping and aggressive behaviour – the experimental condition influenced behaviour through its influence on intensity of sacrifices and identity fusion (see Figure 6).

Figure 6. Lineal mediation model in Study 6 (N = 771 vs. 376 for waves 1 and 2, respectively)



For helping behaviour, condition influenced helping behaviour via 1) identity fusion alone, b = 0.11, 95% CI [0.01, 0.25], 2) intensity of the sacrifice alone, b = 0.57, 95% CI [0.07, 1.05], and 3) intensity of sacrifice and fusion with country serially, b = 0.09, 95% CI [0.003, 0.07]. For aggressive behaviour, condition influenced aggressive behaviour via 1) identity fusion alone, b = -0.07, 95% CI [-0.16, -0.01], 2) intensity of the sacrifice alone, b = -0.36, 95% CI [-0.64, -0.09], and 3) intensity of sacrifice and fusion with country serially, b = -0.06, 95% CI [-0.12, -0.01].
Note. * p < .05; ** p < .01; *** p < .001

General Discussion

In field and online studies, we found that admiration is an emotion that makes individuals feel viscerally connected to their group or convictions, which in turn influences them to engage in costly sacrifices for their comrades or deep-seated beliefs. This effect was robust: it replicated across a variety of research designs (cross-sectional, experimental, and longitudinal), samples (terrorist and non-terrorist), and operationalizations of self-sacrifice (self-report and behaviour). In the prisons of Spain, for instance, the Jihadists we surveyed expressed greater admiration for radical Islamist groups than non-Jihadist Muslims, and such experience was positively related to identity fusion and costly sacrifices for religion in prison.

Five experiments provided additional support for this model and illuminated the conditions under which admiration produces greater identity fusion and costly sacrifices. Together, these findings reveal how admiration may lead people down a path toward extremism, and offer a number of important theoretical and practical implications.

First, admiration for a person or a group produces a chain reaction culminating in individuals making sacrificial behaviours for their ideological convictions. Admiration produces this effect by rendering cognitively salient the sacrifices peers have made for the group – a relationship that has not been documented in previous work. As we consistently showed, the more these sacrifices are perceived as costly, the more people feel fused with their ideological convictions or with their group, and the greater their willingness to fight and die for it. Thus, these results reinforce the notion that admiration is an other-praising emotion connected to virtue (Algoe an& Haidt, 2008; Onu et al., 2016) and explain the psychological process through which admiration galvanises individuals to make self-sacrifices.

Second, admiration produces changes in the self. This finding is important given that prior research has been relatively quiet on the role of emotions in promoting identity fusion. While research has indicated that, in particular, sharing intense dysphoric experiences with other group members causes identity fusion (Jong et al., 2015; Páez et al., 2015; Segal et al., 2018), we show that there are also positive emotions that can enhance the sense of oneness with the group. Importantly, such experience does not necessarily need to be shared with other groups members to produce its effects. Furthermore, given that admiration has been a topic traditionally examined through the lens of positive psychology, it has been widely assumed that admiration *invariably* produces positive effects on intergroup relationships (Onu et al., 2016). The present research challenges this assumption by showing that admiration has a darker side, and can make individuals more prone to defending the group they cherish at all costs,

including through violence. This suggests that admiration, commonly a catalyst for good deeds, can also inspire malevolent actions.

Third, we establish the multiple boundary conditions under which admiration increases identity fusion and greater willingness to fight and die for a group or a set of convictions. For instance, previous work has shown that strongly fused individuals are just as willing to self-sacrifice for one ingroup member as they are for five ingroup members (Swann, Gómez et al., 2014). The present research builds on these findings with evidence that the effect of admiration on the intensity of perceived sacrifices, identity fusion, or willingness to fight and die for the group is not moderated by the number of ingroup members being admired.

In parallel, we also demonstrate that admiration for both ingroup and outgroup members is positively related to people's willingness to fight and die for their group through perceived sacrifices and identity fusion. Although the effect was significantly stronger when admiring an ingroup member, these results suggest that admiration fuels sacrifices even if it occurs in relation to a group that one does not belong to – a finding that warrants further investigation. Additionally, we show that the factors giving rise to admiration are important in this process. Admiring someone for their group sacrifices yields stronger levels of identity fusion and willingness to fight and die for the group than when admiration is related to the subject's personal qualities. This finding is consistent with research indicating that admiration produced by moral actions (sacrifices) — also referred to as moral elevation — versus non-moral actions (skills and talents) is more strongly associated with prosocial tendencies (Algoe & Haidt, 2009).

Fourth, admiration for an ingroup member produces measurable behavioural changes for as long as one month after experiencing this emotion. In this sense, our results are consistent with research showing that moral elevation produces altruistic behaviour (Aquino et al., 2010; Schnall et al., 2010). However, our findings extend prior work in several ways. For starters, we

demonstrate the psychological mechanisms through which this phenomenon occurs, namely (1) perceived intensity of sacrifices and (2) identity fusion. Next, by demonstrating that admiration for a specific ingroup member increases prosocial behaviour – and decreases aggression – toward other ingroup members, we show that the emotional effect generalises to other group members as well. Finally, the results indicate that admiration has relatively long-term consequences that go beyond momentarily priming in an experimental context.

Collectively, our findings support the notion that admiration plays a key role in fostering the emulation of ideals embodied by models of virtue (Algoe & Haidt, 2009; Onu et al., 2016). It does so by increasing the salience of other people's sacrifices, which then produces important changes in identity by aligning the personal self with the collective self. This in turn makes individuals prone to engage in personal costly sacrifices for their ingroup or their convictions.

The preceding results yield important practical implications relevant to preventing and countering violent extremism. For instance, our research suggests that expressing admiration for individuals or groups endorsing violence might be a precursor to radicalization. This finding is particularly relevant where online radicalization is accelerated by propaganda videos conspicuously designed to attract new followers with a narrative intended to produce sympathy and admiration for their members.

Ensuring that individuals, especially youths, have access to prosocial groups (such as sports teams, youth centres, and volunteer organizations) and inspiring role models might be an effective strategy to steer people away from radical organizations. Although this may seem obvious to seasoned practitioners, the signs of radicalization are not always clear to parents, teachers, and community members. Making information more widely available could lead people to seek professional help in a timely fashion (Bélanger, 2017).

Additionally, our research may be useful for designing deradicalization strategies in which the primary goal is to redirect violent extremist offenders away from engaging in or supporting violent activities (Kruglanski et al., 2014; Webber et al., 2018). Our results indicate that jihadists' admiration for terrorist groups is one important predictor of their willingness to engage in extreme behaviour in prison. Given that admiration is in part a moral emotion, one potential strategy to undermine this feeling would be to subvert the radical group's moral stature, for example, by highlighting the inconsistencies between its "moral code" (establishing a caliphate for Muslims) and its behaviours (mass executions of Muslims). Case studies have shown that a large proportion of defectors arrive at these disillusionments on their own (Bjørgo, 2011; Bjørgo and Horgan, 2009). However, additional research is needed to examine how this strategy can be implemented systematically to catalyse attitudinal change in violent extremist offenders.

Notwithstanding the role of admiration in violent extremism, our research should be understood in a broader context, as it is well documented that admiration also produces positive effects on intergroup relationships (Onu et al., 2016; Schnall et al., 2010; Vyver & Abrams, 2015). Moreover, admiring individuals or groups making costly sacrifices for the common good can be an important vector of long-lasting social change through non-violence. For example, in 2019, the environmental activist Greta Thunberg galvanised young climate activists worldwide by skipping school on Fridays to protest in front of the Swedish Parliament for stricter environmental regulations. Her sacrifices culminated in what the editors of *Time* magazine said was the creation of "a global attitudinal shift, transforming millions of vague, middle-of-the-night anxieties into a worldwide movement calling for urgent change" (Time, 2019). More recently, people around the world have cheered healthcare professionals heading to work as they risk their lives to save those infected during the COVID-19 pandemic. It could well be that admiration for these modern-day heroes is motivating others to make personal

sacrifices (such as self-isolation and social distancing) to "flatten the curve" and help their country recover.

Our studies have several limitations as well. First, we have tested our model in samples of Islamic terrorists and general population. Although there is a solid rationale to suppose that the same processes would apply to other groups such as social activists or non-religious terrorists, future studies should test the validity of our theoretical model with different samples. Second, classic socio-psychological research has consistently shown that the probability of success of social influence attempts increases as the number of people who hold an opinion or engage in a certain behaviour augment (Cialdini & Griskevicius, 2010; Cialdini & Sagarin, 2005). Contrarily to that, our results show that the effects of admiration on identity fusion and pro-group behaviour do not vary as a function of the number of individuals toward whom admiration is felt. Identity fusion is based on the experience of strong relational ties with other group members and we believe that this fact may explain our results, but further research is needed to find out if that is the case. Third, several studies suggest that positive mood might drive individuals to engage in prosocial behaviours (e.g., Baron, 1997; North et al., 2004; Salowey et al., 1991). Feeling admiration is a mood enhancer (Algoe & Haidt, 2009; Onu et al., 2016) and, although a high number of studies have demonstrated that the effects of admiration cannot be reduced to those of other positive emotional states (Pohling & Diesner, 2016), future studies should include measures of mood and other positive emotions and control for these variables.

Conclusions

Fighting violent extremism is a collective effort, which requires understanding the motives that drive some individuals to fight, die, and even kill for a group or a set of convictions. Here, we show that admiration enables collective engagement by bringing to mind the sacrifices others have made for the group, which in turn produces a union between the

personal and the social self. The costlier the sacrifices, the greater the identity fusion with other group members, even when the admiration stems from the sacrifices of outgroup members. What's more, intensified identity fusion due to admiration of an ingroup member increases prosocial behaviour toward other ingroups members. Taken together, the present research suggests that admiration is an emotion intimately related to self-sacrifice, which has important implications for group processes, including social transmission and cooperation.

CHAPTER 4

Identity Fusion Predicts Violent Pro-group Behaviour when It Is Morally Justifiable

Reference

Chinchilla, J., Vázquez, A., & Gómez, Á. (2021). Identity fusion predicts violent behavior when it is morally justifiable. *The Journal of Social Psychology*. https://doi.org/10.1080/00224545.2021.1948813. JIF, 2020: 2.71 (Q2).

Abstract

Identity fusion is a visceral feeling of oneness with a group that predicts extreme pro-group behaviours. However, the effects of fusion depending on the nature of such behaviours -violent versus nonviolent- and the factors that may incline strongly fused individuals to display them still remain unexplored. To fill this gap, we performed two correlational studies in which we examined whether moral beliefs regarding the unjustifiability of violence moderate the relationship between fusion with the family (Study 1), or with the country (Study 2), and willingness to engage in violent and nonviolent pro-group acts. Results showed that strongly fused participants were more willing to act violently than weakly fused participants, but only when their beliefs in the moral unjustifiability of violence were weak. In contrast, their willingness to engage in nonviolent acts was not influenced by moral beliefs.

Keywords: Identity fusion, self-sacrificial behaviour, violence, morality

Introduction

Identity fusion is a visceral feeling of oneness with a group that predicts extreme pro-group behaviours. Recent research has demonstrated that this feeling of extraordinary connection can also be established to another individual, an animal, an object, or an activity (Gómez et al., 2020). Although dozens of studies conducted in five continents and in different contexts have consistently shown that identity fusion predicts costly sacrifices and willingness to fight and die for the group among others, and preliminary research has used identity fusion as an approach to explain football violence (Newson, 2017), no empirical research up to date has explored whether the nature of such pro-group behaviour (violent vs. non-violent) affects the predictive character of fusion, and whether beliefs in the moral justifiability of violence moderate the relationship between fusion and pro-group actions. Identifying the variables that determine whether strongly fused individuals pursue the goals of the group through violent versus non-violent activities would provide important insights into the conditions under which identity fusion has negative or positive social consequences (Swann & Buhrmester, 2015), and that is the main aim of this research.

Research on identity fusion originated in the aftermath of the 9/11 attacks in New York, and the Madrid train bombings, as an attempt to explain why some terrorists and members of other violent groups engage in extreme self-sacrificial behaviours (Swann et al., 2009). Identity fusion theory postulates that this type of acts might result from a psychological state in which individuals feel so deeply entrenched into the group and emotionally committed to it that they are willing to promote its interests and that of its members even when that comes at a high cost to the self (Gómez et al., 2020; Swann et al., 2014). Consequently, a considerable amount of research has included measures of willingness to fight and die and to engage in costly pro-group sacrifices and has found that fusion is positively related to them (e.g., Gómez, Brooks et al., 2011; Gómez, Morales et al., 2011; Gómez et al., 2017; Swann, Gómez, Huici et al., 2010;

Swann et al., 2009). Other research has also shown that the tendency of strongly fused individuals to put the group ahead of themselves increases in circumstances of group threat (Sheikh et al., 2016) and that, in some cases, this might boost their aggressive inclinations as well (Vázquez et al., 2020).

The focus of this area of research in the study of terrorism has given rise to some doubts about the nature of the relationship between fusion and violence; and some authors question themselves whether we should always expect a positive relationship between both variables or not (Kiper & Sosis, 2018). Albeit such doubts are easily understandable, identity fusion theory only postulates that fusion drives people to put their behaviour at the service of the well-being of the group and to pay important personal costs in order to do that. The specific form that these sacrifices take (violent vs. non-violent) is supposed to vary as a function of ideological factors, including group norms and personal moral beliefs (Gómez et al. 2020; Whitehouse, 2018). In this vein, previous research suggests that one of the causes of fusion, as it is transformative experiences shared with other group members, might constitute a risk factor that increases the likelihood of individuals lending support, or engaging in, violent behaviours when they share an ideology that considers violence as legitimate (Kavanagh et al., 2020). However, given the youth of the theory, the impact of such ideological factors on fused individuals' behaviour still remains unexplored. We wanted to address this gap through two studies that explore whether personal beliefs regarding the moral justifiability of violence moderate the relationship between identity fusion and pro-group violence without influencing the relationship between fusion and extreme non-violent pro-group acts.

Identity Fusion, Threats, Ideology, and Extreme Pro-group Behaviours

Identity fusion is a type of psychological bond with a group that takes place when the borders between the personal identity - the characteristics of individuals that make them unique - and social identity - the characteristics of individuals that derive from their group membership

- become porous or permeable. As a result of that, both identities stay simultaneously active and synergistically motivate pro-group behaviours. Moreover, fused persons experience deep feelings of personal agency - capacity to initiate and control intentional behaviours - that they put at the service of the group's interests. Additionally, fused persons feel attracted to other group members because of their personal characteristics as well as of their group membership and perceive them as if they were brothers or sisters. Lastly, group membership gives meaning to the personal and social identities of fused individuals what causes that once individuals become fused with a group, they tend to remain fused (Swann et al., 2012).

The studies on the behavioural effects of fusion have consistently found that this phenomenological experience of oneness with a group gives rise to extreme behaviours. For example, identity fusion predicts willingness to (1) fight and die for the group (Bortolini et al., 2018; Gómez, Brooks et al., 2011; Gómez, Morales et al., 2011; Gómez et al., 2017; Swann, Gómez, Huici et al., 2010; Vázquez et al., 2017); (2) die in order to kill terrorists who threaten the group (Swann, Gómez, Dovidio et al., 2010); (3) volunteer for armed combat (Whitehouse et al., 2014); (4) die for one or more ingroup members in several versions of the trolley dilemma (Gómez, Brooks et al., 2011; Swann et al., 2014); or (5) engage in irreversible surgical change of the primary sexual characteristics in individuals suffering from gender dysphoria (Swann et al., 2015). In addition to that, some researchers have thought about the contextual factors that impact the relationship between fusion and the tendency to engage in extreme pro-group behaviours and have concluded that circumstances in which the personal or the group identity are under threat might make this tendency stronger (e.g., Gómez et al., 2011; Sheik et al., 2016; Swann et al., 2009; Vázquez et al., 2020).

The assumption that group threats increase the disposition of highly fused individuals to go to extremes to protect the group has been linked to evolutionary accounts that see human prehistory as a time penetrated by frequent intergroup conflicts and hostilities: a time in which

dying by the hands of opposing group members, or as a result of a battle for scarce resources, were no extraordinary things; and in which participation in highly cohesive groups use to draw the line between survival and death (Atran, 2020; Whitehouse et al., 2017; Whitehouse, 2018). According to this view, identity fusion works as the central node of a psychological machinery that allows for the emergence of highly cohesive coalitions of individuals who are so strongly committed to the group and its members that they are willing to perform extraordinary sacrifices to protect them (Atran, 2020; Swann et al. 2012). Strongly fused persons feel one with the group and perceive other group members as sisters and brothers and, as a result of that, they experience group threats as personal threats which, in turn, leads them to defend and protect the group with the same passion and intensity with which non-fused persons typically try to protect themselves (Gómez, Brooks et al., 2011; Swann et al., 2009, 2014; Whitehouse, 2018).

There is a solid base of empirical evidence in support of the previous view. For instance, Sheikh et al. (2016), performed a study showing that willingness to engage in costly sacrifices for a value was stronger among devoted actors - participants who were fused with a group that shared this value and who saw the value as sacred or non-negotiable - than among non-devoted actors, although the difference between both groups only emerged after participants were reminded that the value was under threat and did not appear under non-threatening circumstances. And Newson et al. (2018) found that Brazilian fused football hooligans were more willing to fight and die for their fellows than fused mainstream fans, which, as the authors suggest, could be due to the fact that hooligans tend to be exposed to more frequent and serious threats than mainstream fans.

Convergently, other studies have demonstrated that threats to the group may also strengthen the aggressive inclinations targeted at the enemies or rivals of the group among the strongly fused. In particular, Vázquez et al. (2020) conducted several studies using a

videogame that allowed participants to annihilate ingroup and outgroup symbols to measure their aggressive inclinations and they found that devoted actors showed more aggressive inclinations than non-devoted actors, but only when the idea that the outgroup constituted a threat to the ingroup had being primed. Complementarily, Fredman et al. (2017) conducted a longitudinal study before, and soon after, the start of the Palestinian Stabbing Intifada to see whether identity fusion impacted support of retaliatory policies under different levels of threat. They discovered that fusion with Judaism was positively related to support of retaliation, and that the link between both variables was stronger after the Intifada began. Finally, Newson et al. (2018) found that highly fused hooligans had engaged in more violence directed at rival fans than highly fused mainstream football fans, and that both groups did not differ in their engagement in violence against non-rivals. Among weakly fused participants, there were no differences as a function of group membership.

The former studies strongly support the view that group threats act as a catalyst of extreme pro-group behaviours among the fused and, at the same time, they show that they might not be a sufficient condition for them to engage in violence. Specifically, the highest rates of past engagement in violence targeting the rivals of the team found among football hooligans by Newson et al. (2018) might have been due to the fact that hooligans are exposed to higher levels of threat than mainstream football fans, to the ideological differences that exist between the groups, or to some combination of both factors. Also, albeit Fredman et at. (2017) found an increase in endorsement of retaliatory policies after the beginning of the Intifada among participants who were fused with Judaism, this effect did not replicate among participants fused with Israel, which, as the authors pointed out, also hints to the possibility that beliefs and ideologies play a key role in the determination of the type of behaviours (violent vs. non-violent) in which fused people decide to engage. Nevertheless, the research conducted up to date has not systematically examined the impact of such ideological factors. Here, we

undertook the exploration of one of them, specifically, beliefs regarding the moral justifiability of violence.

Moral Beliefs, Identity Fusion, and Extreme Violent and Non-violent Pro-group Behaviours

Three different lines of research suggest that moral beliefs may moderate the effects of identity fusion. First, Swann et al. (2014) conducted several studies with two different versions of the trolley dilemma in which participants were asked to choose between sacrificing their own lives to save the lives of one or several ingroup members or sacrificing the lives of several ingroup members to save their own lives. Results of these studies manifested (1) that, although all participants considered that self-sacrifice was the morally correct course of action, only the strongly fused chose to sacrifice themselves; (2) that their self-sacrifice was motivated by visceral commitment to the group and feelings of personal distress in response to the plight of the group members rather than by a lack of concern with their self-preservation; (3) that they were more willing to self-sacrifice under time pressure; and (4) that they were insensible to utilitarian considerations regarding the number of lives that could be saved by their actions. Conjointly considered, these results indicate that the relationship between identity fusion and self-sacrificial acts may be due to a deontic imperative that propels individuals to "renounce whatever is necessary to protect the group or its members" and that moral motives play a remarkable role in the determination of fused individuals' pro-group behaviours.

Second, drawing from previous research in the area of motivation, such as the research by Higgins (1997; Higgins et al., 1994) which posits a dual-regulation system with a prevention focus (based on needs for security and focused on negative end-states), and a promotion focus (based on needs for achievement and focused on positive end-states), Janoff-Bulman et al. (2009) presented several studies pointing out that there are two systems of moral regulation which differ with regard to their motivational force: the proscriptive and the prescriptive moral

regulation systems. The *proscriptive moral system* focuses on what we should not do and involves avoidance motives, overcoming our negative desires, and restraining our motivation to do something bad; whereas the *prescriptive moral system* focuses on what we should do and entails approach motives, overcoming our inertia, and activating our motivation to do something good. In other words, proscriptive morality implies the inhibition of harmful behaviours, and prescriptive morality implies engagement in behaviours that help others by relieving their suffering or advancing their well-being. Because "bad is psychologically stronger than good", the avoidance-based focus on negative (vs. positive) outcomes in proscriptive morality gives rise to a motivational asymmetry between both systems in which the proscriptive system is harsher, stricter, and more demanding than the prescriptive system. The distinction established by Janoff-Bulman et al. (2009) between both systems may help us to understand how fusion and pro-group violence relate.

Strongly fused people feel morally compelled to perform self-sacrifices to protect the group and advance its well-being (Swann et al., 2014), which is a type of motivation that can be framed as pertaining to the prescriptive moral system. If they hold the belief that violence on behalf of groups is morally justifiable, the prescriptive tendency to protect their group could drive them to act more violently against threatening individuals or groups than their weakly fused counterparts. However, if strongly fused persons do not hold the belief that violence on behalf of groups is morally justifiable, inasmuch as the proscriptive moral system is harsher than the prescriptive one, the moral condemnation of violence should prevent them from engaging in violent pro-group behaviours, without influencing their disposition to engage in extreme non-violent actions. That is, not holding moral beliefs that blatantly oppose the use of violence should be a prerequisite for the fused to engage in pro-group violence.

Third, Bairak (2019) explored the impact of the moral foundations postulated by Curry (Curry, 2016; Curry et al., 2019) on the relationship between group identification and ingroup

favouritism. She found out that group identification was positively related to ingroup favouritism among participants who did not endorse the foundation of reciprocity, whereas both variables were unrelated among participants who endorsed it. It is important to note, however, that there are substantial differences between endorsement of reciprocity and endorsement of the idea that violence is morally justifiable. Likewise, identity fusion and group identification are two different kinds of psychological alignment with the group, and there are many processes, and mechanisms, in which they differ (for a brief summary of the differences and similarities between both constructs see Gómez et al., 2020). Despite these differences with respect to Bairak's (2019) design, we also expect that moral beliefs about violence would moderate the relationship between fusion and engagement in extreme violent acts. In the present research, this hypothesis was put to the test.

Overview of the Studies

We performed two correlational studies. Study 1 examined whether beliefs regarding the moral justifiability of violence on behalf of groups moderate the relationship between fusion with family and willingness to engage in extreme violent acts in order to protect the family's stability and continuity. Additionally, Study 1 tested whether such beliefs do not impact the relationship between fusion with family and willingness to engage in extreme non-violent acts on its behalf.

In Study 2, we wanted to replicate and generalize the results of Study 1 to a different group, focusing on the country rather than on the family.

We did not determine sample size a priori. Both studies where open for a week and then were closed. Participants in both studies were recruited using a snowball procedure wherein university students from an open university invited their acquaintances to volunteer. The materials of the two studies and the data that support the findings are openly available in "Open Science Framework" at https://doi.org/10.17605/OSF.IO/HM5YC

Study 1

Study 1 was designed to test whether beliefs regarding the moral justifiability of violence moderate the relationship between fusion with the family and willingness to engage in extreme non-violent and violent behaviours when the stability and continuity of the family is under threat. On the one hand, we expected that strongly fused participants would be more willing to engage in *extreme non-violent acts* for their family than weakly fused participants regardless of their moral beliefs. On the other hand, we anticipated that strongly fused participants would be more willing to engage in *extreme violent acts*, but only when they believe in the moral justifiability of violence.

Method

Participants

Four hundred and seventy-eight participants volunteered for the study (95.6% Spaniards, 62.1% women, M_{age} = 37.68, SD = 12.91), and it was administered on-line, via Qualtrics. A sensitivity analysis conducted with G*Power (Faul et al. 2007) after participants' recruitment revealed that this sample (N = 478) had 80% power to detect a minimum effect size of f^2 = 0.02 in a multiple regression with three predictors, asuming an alpha significance criterion of 0.05.

Procedure

First, participants responded to a 3-item measure of *moral beliefs in the justifiability of violence* (e.g., "Because of moral objections, I would never justify the use of violence to protect the continuity and stability of a group", "I would find it very difficult to be friends with someone who believes it is morally justifiable to kill in order to protect the stability and continuity of a group", α = .73) adapted from Zaal et al. (2011). Next, participants answered to the 7-item verbal measure of *fusion with family* (e.g., "I am one with my family", "I make

my family strong", α = .80) by Gómez, Brooks et al. (2011). In this study, and also in Study 2, we preferred to use the verbal measure of fusion instead of the original pictorial measure (Swann et al., 2009) or the Dynamic Identity Fusion Index (DIFI, Jiménez et al., 2016) because of the explicit recommendation of the authors to use the verbal option when it is possible (see also Gómez et al., 2020). Finally, we asked participants to imagine that the stability and continuity of their family is under a heavy threat, and complete a 5-item measure of *willingness to engage in extreme non-violent behaviour* (e.g., "I would be willing to openly express the idea that the threat should stop, even if I were at risk of being rejected by people important to me", "I would be willing to die for organizing strikes and demonstrations against this threat", α = .82)[3], followed by a 5-item measure of *willingness to engage in violent behaviour* (e.g., "I would be willing to use violence against other people to end this threat", "I would be willing to kill to end this threat", α = .95), developed for the study.

We performed a principal axis factor analysis on the items of the last two scales with oblique rotation (direct oblimin). Results manifested that the items loaded on two different

[3] We planned the two studies included in the paper simultaneously and considered the possibility of using different scales to measure extreme non-violent acts, but a few weeks before we started collecting the data for Study 1 two workers at the landfill of Zaldivar (in the Basque Country) were buried alive by a rubbish avalanche, and the management of the rescue and recovery tasks by local authorities was so poor that it caused a wave of acts of protest and social mobilization, many of which were organized by the own families of the workers. Given that these events were widely reported by the Spanish media and all the population was presumably aware that there are circumstances in which it might be reasonable to engage in political acts to protect the interests and integrity of the family, we changed our mind and decided to use the same scale in both studies to increase the comparability of the results (please, note that the sample was mostly composed by Spaniards). None of the participants in Study 1 with whom we talked after the study told to us that any of the items was unrealistic or unbelievable.

factors, with each item loading onto the expected factor, and factors loadings ranging from .673 to .966 (see Supplementary Materials, SM). Responses to the measures used in all studies ranged from 1 (*completely disagree*) to 7 (*completely agree*). Higher scores indicate lower beliefs in the justifiability of violence and higher levels of fusion with the family and willingness to engage in extreme non-violent and violent behaviour, respectively.

Results

Table 5 shows the means, standard deviations, and Pearson correlations of the variables. Fusion with the family was positively related to all the variables. Moral beliefs were positively related to willingness to engage in extreme non-violent acts and negatively related to willingness to engage in violent acts. Lastly, willingness to engage in extreme non-violent acts was positively related to willingness to engage in violent acts.

Table 5. Means, standard deviations, and correlations between variables (Studies 1 and 2)

		M	*SD*	1	2	3
Study 1 (Fusion with the Family)	1. Moral Beliefs	5.16	1.35	-		
	2. Identity Fusion	4.84	1.15	.11*	-	
	3. Non-violent Behaviour	4.54	1.33	.12**	.27**	-
	4. Violent Behaviour	3.18	1.86	-.32**	.14**	.32**
Study 2 (Fusion with the Country)	1. Moral Beliefs	5.67	1.42	-		
	2. Fusion with Country	3.30	1.45	-.10	-	
	3. Non-violent Behaviour	3.26	1.19	-.04	.23**	-
	4. Violent Behaviour	1.86	1.27	-.37**	.24**	.34**

Note: * $p < .05$; ** $p < .01$

Given that our dependent measures were nested within participants, we first performed a two-level mixed model analysis with random intercept, including identity fusion (mean centred), moral beliefs (mean centred), type of behaviour (0 = violent; 1 = non-violent), and the two-way and three-way interactions as fixed effects predictors, and participants' answers to the scales of willingness to engage in extreme non-violent and violent behaviours as

dependent variable. We fitted a variance components covariance structure for the random effect and used maximum likelihood to estimate the parameters in the analysis.

Table 6 shows the parameters estimates of the effects included in the model. Results showed that the three-way interaction between identity fusion, moral beliefs, and type of behaviour was not significant, $p = .211$. However, the expected two-way interaction between fusion and moral beliefs, and between moral beliefs and type of behaviour were significant, *ps* $\leq .002$, indicating that the effect of identity fusion on the willingness of participants to engage in pro-family behaviours varied as a function of their moral beliefs, and that the effect of moral beliefs depended on the type of behaviour. All the main effects were significant too, $p < .001$.

Table 6. Parameter estimates for the effects included in the model (Study 1; fusion with the family)

Fixed Effects	***B***	***SE***	***t***	***p***	**95% CI**
Intercept	3.20	0.07	48.10 (871.01)	< .001	3.07, 3.33
Fusion with the Family	0.29	0.06	5.01 (871.01)	< .001	0.18, 0.41
Moral Beliefs	-0.63	0.05	-13.05(871.01)	< .001	-0.72, -0.53
Type of Behaviour	1.35	0.08	17.27 (478)	< .001	1.19, 1,50
Fusion X Moral Beliefs	-0.13	0.04	-3.12 (871.01)	.002	-0.21, -0.05
Fusion X Type of Behaviour	0.02	0.07	0.35 (478)	.722	-0.11, 0.16
Moral Beliefs X Type of Behaviour	0.62	0.06	11.09 (478)	< .001	0.51, 0.73
Fusion X Moral Beliefs X Type of Behaviour	0.06	0.05	1.25 (478)	.211	-0.03, 0.15
Variance Components	**Estimate**	***SE***	***Z***	***p***	**95% CI**
Random Intercept Variance	0.66	0.10	6.56	< .001	0.49, 0.88
Residual Variance	1.45	0.10	15.46	< .001	1.27, 1.64

Note: *ts* degrees of freedom are included between parentheses.

We conducted further separate multiple regression analyses on willingness to engage in extreme non-violent and violent acts. For the analyses of both dependent variables, moral

beliefs (mean centred), fusion (mean centred), and the two-way interaction term were entered simultaneously as predictors into the analyses.

Extreme Non-violent Behaviour

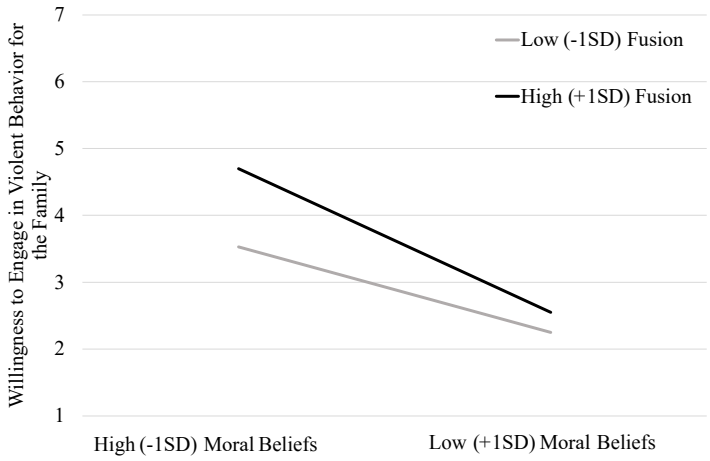
The effect of the overall model was significant, $F(3,474) = 13.96$, $p < .001$, $R^2 = .08$, $f^2 = 0.09$. Results revealed a significant main effect of fusion, $B = 0.31$, $t(474) = 6.18$, $p < .001$, 95% CI [0.21, 0.42], such that willingness to engage in extreme non-violent acts increased as fusion went stronger. The main effect of moral beliefs was not significant, $B = 0.00$, $t(226) = -0.06$, $p = .954$, 95% CI [-0.08, 0.08], neither was the interaction between fusion and moral beliefs, $B = -0.07$, $t(226) = -1.88$, $p = .07$, 95% CI [-0.14, 0.00].

Violent Behaviour

The effect of the overall model was significant, $F(3,474) = 53.77$, $p < .001$, $R^2 = .25$, $f^2 = 0.33$. The main effects of fusion, $B = 0.29$, $t(474) = 4.50$, $p < .001$, 95% CI [0.16, 0.42], and of moral beliefs, $B = -0.63$, $t(474) = -11.71$, $p < .001$, 95% CI [-0.73, -0.52], were significant as well, such that willingness to engage in extreme violent acts increased as fusion and moral beliefs went stronger.

Results showed a significant interaction between identity fusion and moral beliefs too, $B = -0.13$, $t(474) = -2.80$, $p = .005$, 95% CI [-0.22, -0.04]. Simple slope analyses revealed that fusion with the family was positively related to willingness to engage in extreme violent behaviour among participants who agreed with the idea that violence is morally justifiable, $B = 0.47$, $t(474) = 5.18$, $p < .001$, 95% CI [0.29, 0.65], but not among participants who did not thought that violence is morally justifiable, $B = 0.11$, $t(474) = 1.27$, $p = .204$, 95% CI [-0.07, 0.29], see Figure 7.

Figure 7. Willingness to engage in violent behaviour for the family as a function of identity fusion and moral beliefs regarding the justifiability of violence



Discussion

Study 1 indicated that, among participants who held strong beliefs about the moral justifiability of violence, willingness to engage in violence on behalf of the family increased as increased the level of fusion. By contrast, among participants who held weak beliefs about the moral justifiability of violence, willingness to engage in violence did not vary as a function of fusion with family. At the same time, Study 1 revealed that the moderating role of beliefs in the moral justifiability of violence did not extend to willingness to engage in extreme non-violent behaviours, which, as expected, only increased as increased the degree of fusion.

Study 2

Study 2 was designed to test whether the results found in Study 1 replicate when considering a different focal group: the country. As in the previous study, we expected that strongly fused participants would be more willing to engage in *extreme non-violent acts* for their country than weakly fused participants regardless of their moral beliefs. Additionally, we

hypothesized that strongly fused participants would be more willing to engage in *extreme violent acts* than weakly fused participants if participants believed that violence is morally justifiable, but there would be no differences if participants did not believe that violence is morally justifiable.

Method

Participants

Two hundred and thirty participants volunteered for the study (97.8% Spaniards, 63.9% women, M_{age} = 40.10, *SD* = 13.14), and it was conducted on-line, via Qualtrics. A sensitivity analysis conducted with G*Power (Faul et al. 2007) after participants' recruitment revealed that this sample (n = 230) had 80% power to detect a minimum effect size of f^2 = 0.04 in a multiple regression with three predictors, asuming an alpha significance criterion of 0.05.

Procedure

Moral beliefs, *identity fusion*, *willingness to engage in extreme non-violent behaviour*, and *willingness to engage in violent behaviour* were measured with the same procedure and scales used in Study 1 adapted to the country, αs = .70, .89, .80, and .95, respectively. As in the previous study, we asked participants to imagine that the stability and continuity of the country is under threat before measuring their behavioural inclinations. We conducted a principal axis factor analysis on the items of the last two scales with oblique rotation (direct oblimin). Results of the analysis showed that the items loaded on two different factors, with each item loading onto the expected factor, and factors loadings ranging from .558 to .975 (see SM).

Results

Table 5 shows the means, standard deviations, and bivariate correlations between the variables. Fusion with the country was positively related to willingness to engage in extreme

non-violent and violent behaviours, willingness to engage in extreme non-violent behaviours was positively related to willingness to engage in violent behaviours, and moral beliefs were negatively related to willingness to engage in violence. The relationships between identity fusion and moral beliefs, and between moral beliefs and extreme non-violent behaviours were not significant.

We performed a two-level mixed model analysis with random intercept, including identity fusion (mean centred), moral beliefs (mean centred), type of behaviour (0 = violent; 1 = non-violent), and the two-way and three-way interactions as fixed effects predictors, and participants' answers to the scales of willingness to engage in extreme non-violent and violent behaviours as dependent variable. We fitted a variance components covariance structure for the random effect and used maximum likelihood to estimate the parameters in the analysis.

Table 7 shows the parameters estimates of the effects included in the model. The two-way interaction between fusion and moral beliefs, and between moral beliefs and type of behaviour were significant, $ps \leq .039$, manifesting that the effect of identity fusion on the willingness of participants to engage in behaviours on behalf of the country varied as a function of their moral beliefs, and that the effect of moral beliefs depended on the type of behaviour. The three-way interaction between fusion, moral beliefs, and type of behaviour was significant too, $p = .029$, showing that the effect of the interaction between identity fusion and moral beliefs differs as a function of the behaviour analysed. All main effects were also significant, $ps \leq .001$.

Table 7. Parameter estimates for the effects included in the model (Study 2; fusion with the country)

Fixed Effects	***B***	***SE***	***t***	***p***	**95% CI**
Intercept	1.84	0.08	24.25 (415.89)	< .001	1.69, 1.99
Fusion with the Country	0.18	0.05	3.42 (415.89)	.001	0.08, 0.28
Moral Beliefs	-0.29	0.05	-5.39 (415.89)	< .001	-0.40, -0.19
Type of Behaviour	1.42	0.09	16.05 (230)	< .001	1.24, 1,59
Fusion X Moral Beliefs	-0.07	0.03	-2.07 (415.89)	.039	-0.13, -0.00
Fusion X Type of Behaviour	0.01	0.06	0.15 (230)	.882	-0.11, 0.13
Moral Beliefs X Type of Behaviour	0.27	0.06	4.35 (230)	< .001	0.15, 0.40
Fusion X Moral Beliefs X Type of Behaviour	0.08	0.04	2.19 (230)	.029	0.01, 0.16
Variance Components	**Estimate**	***SE***	***Z***	***p***	**95% CI**
Random Intercept Variance	0.43	0.09	4.70	< .001	0.28, 0.65
Residual Variance	0.89	0.08	10.72	< .001	0.74, 1.07

Note: *ts* degrees of freedom are included between parentheses.

We also conducted separate multiple regression analyses on willingness to engage in extreme non-violent and violent acts. For each dependent variable, moral beliefs (mean centred), fusion (mean centred), and the two-way interaction term were entered simultaneously as predictors into the analyses.

Extreme Non-violent Behaviour

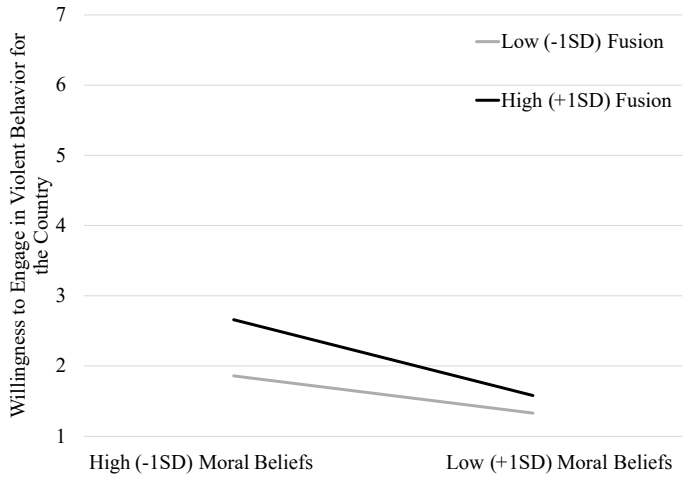
The effect of the overall model was significant, $F(3,226) = 4.33$ $p < .001$, $R^2 = .05$, $f^2 = 0.05$. Results showed as well a main effect of fusion, $B = 0.19$, $t(226) = 3.53$, $p = .001$, 95% CI [0.08, 0.30], indicating that willingness to engage in extreme non-violent acts increased as fusion went stronger. The main effect of moral beliefs was not significant, $B = -0.02$, $t(226) = -0.33$, $p = .738$, 95% CI [-0.13, 0.09], neither was the interaction between fusion and moral beliefs, $B = 0.03$, $t(226) = 0.47$, $p = .640$, 95% CI [-0.05, 0.08].

Violent Behaviour

The effect of the overall model was significant, $F(3,226) = 18.24$, $p < .001$, $R^2 = .20$, $f^2 = .0.25$. The main effects of fusion, $B = 0.18$, $t(226) = 3.38$, $p = .001$, 95% CI [0.07, 0.28], and moral beliefs, $B = -0.29$, $t(226) = -5.40$, $p < .001$, 95% CI [-0.40, -0.19], were significant, indicating that willingness to engage in extreme violent acts increased as fusion and moral beliefs went stronger.

Additionally, there was a significant interaction between identity fusion and moral beliefs, $B = -0.07$, $t(226) = -2.07$, $p = .040$, 95% CI [-0.13, -0.003]. Simple slope analyses showed that fusion with the country was positively related to willingness to engage in extreme violent behaviour among participants who agreed with the idea that violence is morally justifiable, $B = 0.28$, $t(226) = 3.95$, $p < .001$, 95% CI [0.14, 0.41], but not among participants who did not thought that violence is morally justifiable, $B = 0.08$, $t(226) = 1.12$, $p = .263$, 95% CI [-0.06, 0.22] see Figure 8.

Figure 8. Willingness to engage in violent behaviour for the country as a function of identity fusion and moral beliefs in the justifiability of violence



Discussion

Study 2 replicated the pattern of effects founded in Study 1. Specifically, among participants who held strong beliefs about the moral justifiability of violence, there was a positive relationship between identity fusion and willingness to engage in pro-group violence, whereas both variables were unrelated among participants who held weak beliefs about its moral justifiability. Study 2 showed as well that fusion was positively related to willingness to engage in extreme non-violent acts, and that the relationship between both variables was not moderated by moral justifiability of violence. Importantly, Study 2 extended these findings to the country, suggesting that the moderating role of beliefs about the justifiability of violence is important when individuals are fused with extended - large groups in which individuals do not personally know each other - as well as with local groups - small groups of intimate others - (Swann et al., 2012).

General Discussion

The present research explored, for the first time, the moderating role of moral beliefs in the willingness of fused people to engage in pro-group behaviours entailing different degrees of violence. In two studies considering a local (family) or an extended group (country), we found that highly fused participants were more willing to engage in violence to protect their group from a serious threat than weakly fused participants, but only if they believed that pro-group violence is morally justifiable. Among those individuals who did not believe that violence is justifiable, fusion did not predict willingness to use violence.

Our results make several contributions to different domains. At a theoretical level, our studies advance identity fusion research in at least two important ways. First, these studies represent the first systematic exploration of the effects of explicit moral beliefs on fused individuals' disposition to engage in pro-group violence. This being the case, they help solve the doubts about how to interpret the nature of the relationship between fusion and violence

that have been recently expressed within the field (Kiper & Sosis, 2018; Swann & Buhrmester, 2015), showing that whether fused individuals engage or not in pro-group violence when the group is under threat is contingent upon what they believe about the moral justifiability of violence and, by extension, that fusion is neither intrinsically bad nor good. Fused individuals are heavily inclined to protect their group, but they are not necessarily more violent than their non-fused counterparts. Second, our studies complement previous research on the moderators that attenuate fusion effects on extreme self-sacrifices, like knowing that other group members are strongly morally motivated to sacrifice themselves or the degree of certainty with regard to the own level of fusion (Paredes et al., 2018; 2019). These last factors seem to have an indiscriminate or generalized impact on self-sacrifices, whereas our studies identify a new moderator variable that may attenuate, and even cancel out, fusion effects on violence, without influencing its effects on extreme non-violent acts: personal beliefs in the moral justifiability of violence. Complementarily, our studies also advance research on the two systems of moral regulation (Janoff-Bulman et al., 2009), suggesting that the proscriptive moral system is stronger than the prescriptive system even in those individuals who take the protection of the group as a moral mandate, as the fused typically do (Swann et al., 2014).

At a practical level, recent research has discovered that there is a majority of individuals who are fused with their groups among combatants and ISIL terrorists, as well as among persons at risk of radicalization (Gómez et al., 2017; Sheikh et al., 2016; Whitehouse et al., 2017). Related to that, our research highlights new paths that may be followed by program developers and practitioners to tackle the societal and personal problems caused by those people who use terrorism as a means to preserve or promote the well-being of the groups they are fused with. As compared to other strategies suggested before, which are mostly focused on reverting the process of fusion (e.g., Gómez et al., 2019), this new strategy is only focused on the change of moral considerations regarding violence, and, as such, it does not

imperil the sense of personal and social meaning that fused individuals experience as a result of their visceral union with their group.

Of course, our research has several limitations as well. First of all, we did not include manipulations checks to find out if the participants in our studies truly imagined that the stability and continuity of the group were under threat. Including them would have increased our trust in the fact that the participants followed our instructions, but we think that this is a minor concern because all the items of the scales of behavioural intentions explicitly mention the threat and the studies that have been conducted with participants extracted from the same pool have invariably shown that they are very sensitive to verbal instructions that depict the group or its members as being under different levels of threat (e. g., Sheikh et al., 2016; Vázquez et al., 2020). In addition to that, our studies follow a correlational approach and, as such, they do noy allow to directly establish causality. Undoubtedly, obtaining additional experimental evidence would be advisable, but moral beliefs and convictions are at the nucleus of our identity (Aquino & Reed, 2002); what drives people to defend them in different ways when they face disconfirmatory information and make them extraordinarily resistant at social influence attempts (e.g., Aramovich et al., 2011; Skitka, 2010). On account of that, manipulating moral beliefs regarding the moral unjustifiability of violence would be a really difficult venture.

Our studies highlight some future avenues and lines of research too. Some studies could explore the effects of moral beliefs regarding the moral justifiability of violence on behalf of the specific group to which individuals are fused and compare them to the effects of beliefs regarding violence on behalf of groups in general. Several studies have shown that the relationship between moral beliefs and behaviours is stronger when we measure such beliefs using scales that tap specifically onto the behaviours of interest than when we use scales that do not tap onto them (e.g., Maio, 2016). In consonance to that, we expect that the effects of

moral beliefs referring to the focal group would be stronger than the effects of general beliefs that we have found in our research. Other studies could examine the processes that are behind the development of proscriptive beliefs against violence in fused people. As Wisneski and Skitka (2017), we suspect that some emotions, like disgust or anger, may play an important role in this process of moralization. Research could also examine whether the effect of moral beliefs that we found in our studies extends to other ideological factors, like the group norms or its narratives (Elnakouri et al., 2018; Ginges & Shakleford, 2018; Gómez et al., 2020). In this respect, it has been consistently found that there are notable differences between groups with regard to their moral norms and shared beliefs about the nature of morality, the group, and intergroup relations (Fiske & Rai, 2015; Louis et al., 2018). Besides of that, personal moral convictions are not individually forged. They are developed and shaped through the exposure to the ideas of significant others and can stem from group values and narratives that are gradually internalized (Ellemers, 2017; Kruglanski et al., 2019). Because fused individuals are not an exception to this general rule, we believe that the narratives and norms of the group with which individuals fuse are a key determinant of their predisposition to engage in violence on its behalf. The studies by Fredman et al. (2017) and Newson et al., (2018) already point to this possibility. Finally, it could be interesting to delve into the boundary conditions under which moral proscriptive beliefs exert their effects. Consistently with Bandura's moral disengagement theory (Bandura, 1990), we consider that the impact of this kind of beliefs may be selective rather than universal, and that fused people who believe that violence is unjustifiable will not refrain themselves from acting aggressively if they dehumanize the other, disregard the harm caused by their acts, or are able to displace feelings of responsibility.

Conclusions

Identity fusion theory was developed around a decade ago to capture the psychological processes and mechanisms that explain extreme pro-group behaviours. Since

then, several dozens of studies have shown that this visceral feeling of oneness with the group can drive individuals to engage in extreme acts (like willingness to fight and die or displaying costly self-sacrifices) on its behalf, which has generated doubts about how to interpret the nature of the relationship between identity fusion and pro-group violence (Kiper & Sosis, 2018; Swann & Buhrmester, 2015). The existence of an association between both variables has never been established by the theory, which assumes that the type of pro-group acts (violent vs. non-violent) in which fused individuals engage are determined by ideological factors, such as personal moral beliefs or convictions (Gómez et al., 2020). Our studies confront this tenet with empirical data for the first time and show that, although fused individuals may be highly motivated to protect the group or its members, they are willing to engage in violence only if they believe that that is morally justifiable, but not if they do not believe so. Identity fusion, therefore, should not be taken as a sufficient condition for the emergence of pro-group violence.

CHAPTER 5

Strongly Fused Individuals Feel Viscerally Responsible to Self-sacrifice

Reference

Chinchilla, J., Vázquez, A., & Gómez, Á. (2022). Strongly fused individuals feel viscerally responsible to self-sacrifice. *British Journal of Social Psychology, 00*, 1-19. https://doi.org/10.1111/bjso.12526. JIF, 2020: 4.69 (Q1).

Abstract

Identity fusion is a visceral feeling of oneness that predicts extreme behaviour on behalf of the target of fusion. We propose that strongly fused individuals are characterized by feelings of visceral responsibility towards such target - unconditional, instinctive, and impulsive drive to care, protect and promote its wellbeing and interests - that motivates them to self-sacrifice. Two studies offered initial support when the target of fusion is an individual or a group (Studies 1a-1b). A final study added causal evidence that strongly fused learning that most ingroup members did not feel visceral responsibility towards the group expressed less willingness to self-sacrifice than those learning that ingroup members display high levels of visceral responsibility (Study 2). These findings offer novel evidence for the mechanisms underlying the effects of fusion on extreme behaviour on behalf of the target of fusion and the attenuation of its consequences.

Keywords: Identity fusion, extreme behaviour, prosocial behaviour, care ethics

Introduction

> If I take on the other's reality as a possibility and begin to feel its reality, I feel also that *I must* act accordingly, that is, I am impelled to act as in my own behalf, but in behalf of the other.
>
> Nel Noddings (2003; cursives are ours).

Over the last decade, identity fusion - a visceral feeling of oneness with a group - has been established as one of the most competent predictors of fighting, killing, and dying for a group. Researchers have shown that the effects of fusion are particularly powerful when strongly fused individuals are physically, psychologically, or emotionally activated, which calls for the fact that strongly fused should display intrinsic motivation or inner strength that automatically motivates them and makes them viscerally responsible to act on behalf of the group. This kind of motivational mechanism has been described by care ethicists. Inspired by their view, we define visceral responsibility as an unconditional, instinctive, and impulsive drive to care, protect, and promote the well-being of the target group and its interests. The main goal of this investigation is to determine whether visceral responsibility is characteristic of strongly fused individuals and if this inner strength is an underlying mechanism to explain the effects of fusion on pro-group behaviour.

Identity Fusion

Identity fusion research originated to unravel why some individuals, such as terrorists or social activists, engage in extreme self-sacrificial acts on behalf of the group (e.g., Gómez et al., 2017, 2020, 2021; Gómez, Brooks et al., 2011; Swann et al., 2009, 2012; Whitehouse, 2018). Identity fusion is a visceral feeling of oneness in which the psychological divide that separates the personal self from the group blurs.

In a state of fusion, the personal self - the characteristics of individuals that make them unique - aligns firmly with the social self - the characteristics of people that derive from their group membership - such that both identities remain simultaneously active and synergistically motivate behaviour. Fused people not only develop a strong allegiance to the group category and the values and goals that it represents - collective ties - but also a strong allegiance to fellow group members - relational ties -, who are seen as if they were brothers and/or sisters, even if they are not personally acquainted with them (for an extensive discussion and empirical distinction, see Gómez et al., 2019). Together, these strong allegiances trigger intense feelings of personal agency - capacity to initiate and control intentional behaviours - that are used to selectively advance the interests of the group. Furthermore, because identity fusion is highly demanding and gives meaning to the personal and social self, once people become fused, they tend to remain fused (Swann et al., 2012). One of the most characteristic demonstrations of the interconnection between the self and the group among strongly fused individuals is that when feelings of fusion are experienced, the relational ties established with the group and its members drive strongly fused individuals to react to threats to the group as a personal threat, and to perceive that personal threats also represent a threat to the group (Gómez, Brooks et al., 2011; Swann et al., 2009), which compels them to act on behalf of the group as if they were acting in their own behalf and to incur in great risks and personal sacrifices for it.

Research on the consequences of identity fusion has systematically shown that this experience of psychological oneness motivates willingness for several types of extreme pro-group actions. Identity fusion predicts, among others, willingness to (1) fight and die for the group and its members (Bortolini et al., 2018; Carnes & Lickel, 2018; Gómez, Brooks et al., 2011; Gómez, Morales et al., 2011; Paredes et al., 2018; Swann, Gómez, Huici et al., 2010; Vázquez, Gómez, Swann et al., 2017); (2) die to kill terrorists who threaten the group in an intergroup version of the trolley dilemma (Swann, Gómez, Dovidio et al., 2010); (3) volunteer

for armed combat (Gómez et al., 2017; Kunst et al., 2018; Whitehouse et al., 2014); (4) die to save the life of one or more group members in several versions of the trolley dilemma (Gómez, Brooks et al., 2011; Swann, Gómez et al., 2014; Swann, Gómez, Dovidio et al., 2011); (5) make costly sacrifices for the values that are central to the group (Sheikh et al., 2016); (6) give up important personal relationships to belong to the group (Swann et al., 2015); and (7) engage in violence to protect the stability and continuity of the group when it is seen as morally justifiable (Chinchilla et al., 2021). Importantly, identity fusion does not only predict intentions, but also actual extreme behaviour. For example, Swann et al. (2015) found that transsexuals who were fused with their cross-gender group were more than twice as likely to have undergone irreversible surgical change of their primary sexual characteristics than non-fused transsexuals two years after the assessment of fusion, what represents, at the same time, an extreme behaviour for the self and to belong to the group which is the target of fusion. Also, Gómez el al. (2017) conducted interviews with imprisoned ISIS terrorists and combatants in the frontline against ISIS and discovered that all of them were fused with their group. And a study with a sample of Libyan revolutionaries (Whitehouse et al., 2014) showed that frontline combatants were more likely to be fused with their battalion that individuals who only provided logistical support and, consequently, were less exposed to physical risks.

In addition to that, other studies have found that fusion motivates behaviours that do not demand a strong personal sacrifice, but that nonetheless benefit the groups towards which it is experienced. For instance, people who are strongly fused are more willing to remain in the group after having been ostracized (Gómez, Morales et al., 2011); deny its wrongdoing (Besta et al., 2014); defend its reputation (Ashokkumar et al., 2019); and write support notes to group members who have been victims of a violent extremist attack (Buhrmester et al., 2014). They also are more inclined to make real monetary donations to needy members (Buhrmester et al., 2014; Swann, Gómez, Huici et al., 2010); maximize the economic advantage of the ingroup

over other groups at their own expense (Buhrmester el al., 2018); and exert actual physical effort to race an avatar representing the group (Swann, Gómez, Huici et al, 2010).

Along with the studies focusing on fusion with social groups, other area of research that has received a considerable amount of attention from scientists concerns interpersonal fusion, that is, identity fusion with specific individuals, like the romantic partner, a sibling, or a twin. The results of this research also confirm that when this feeling of extraordinary connection is experienced at the interpersonal level, it motivates extreme and non-extreme prosocial behaviours. Thus, it has been discovered that fusion with an individual predicts willingness to fight and die (Vázquez et al., 2015) and engage in costly sacrifices (Vázquez, Gómez, Ordoñana et al., 2017) for him/her; and die to spare his or her life in an adapted version of the trolley dilemma (Joo & Park, 2018). Fusion with an individual also predicts, among other non-extreme behaviours, forgiveness after having been disappointed (Vázquez, Gómez, Ordoñana et al., 2017); and the use of more constructive responses to relationship conflict (Walsh & Neff, 2018).

But what makes people so extraordinary willing to protect and defend social groups or individuals when they fuse with them? Three different lines of research show that the effects of fusion are especially intense when strongly fused individuals become physically, psychologically, or emotionally activated and highlight the role played by automatic processes and visceral factors in the consequences of being fused. First, regarding physical activation, Swann, Gómez, Huici et al. (2010) showed that elevating strongly fused participants' autonomic arousal through the practice of different types of physical exercise individually or in group - dodgeball, wind sprints, and Exercycle - amplified their tendency to endorse and engage in pro-group behaviours by increasing their feelings of personal agency. Second, concerning psychological activation, several studies have found that threatening the personal or social identity of strongly fused individuals by providing them with feedback inconsistent

with their views potentiates their willingness to engage in extreme self-sacrifices for the group as compared to circumstances in which none of both identities is under threat (Gómez, Brooks et al., 2011; Swann et al, 2009). Lastly, respecting emotional activation, Swann, Gómez et al. (2014) conducted various studies with different versions of the trolley dilemma, wherein participants were asked to choose between sacrificing their own lives to save the lives of one or more ingroup members or sacrificing several ingroup members to save their own lives. They found that, although all participants thought that self-sacrifice was the morally correct option, only strongly fused participants chose sacrificing themselves; and that their willingness to self-sacrifice increased when reflective control was disrupted and was caused by intense feelings of anxiety and emotional distress in response to the plight of the group members rather than by lack of concern with self-preservation. Taken together, these results hint to the possibility that fused individuals' extreme pro-group acts might be motivated by feelings of visceral responsibility like those described by care ethicists.

Care Ethics, Identity fusion, and Feelings of Visceral Responsibility

Care ethics is an approach to ethics underpinned by an ontology of the person that conceives humans as essentially embedded within relationships (e.g., Gilligan, 1982; Noddings, 2003; Slote, 2007; Tronto, 1993). Advocates of the ethics of care see caring as grounded in feeling and contend that, when we value others and engage with them in caring relationships, our passions and social instincts give rise to automatic feelings of responsibility for their well-being (Tronto, 1993, 1998), causing a motivational shift in which behaviour is largely determined by the needs of the cared for without need for abstract reasoning or justification (Noddings, 2003).

Care ethics emerged in the early 1980s as a critique of the abstract and rationalistic conception of morals embraced by deontologism and the liberal conception of justice (Gilligan, 1982), which are built upon an ontology of the person that sees humans as separate and

autonomous beings and, because of that, are obligation or duty based. Considered from their perspective, moral decision makers figure out what obligations they have to respond to the situation and then act accordingly, even - or specially - when they do not want to do so (e.g., Kant, 1788/2011). Opposing this view, care ethics is built upon a conception of humans as beings relationally involved with others; and, consequently, this approach recognizes the moral role played by visceral factors and social instincts.

Care theorists look to Hume (1739/2000) and other moral sentimentalists as forefathers of their approach and see caring as grounded in feelings, receptivity to the needs of the other, and responsiveness (e.g., Noddings, 2003; Tronto, 1993, 1998). As their predecessors do, care ethicists contend that prosocial behaviours cannot be justified solely in rational terms because allowing or causing harm is not necessarily against reason. They also criticize that obligation-based ethics cannot accommodate extreme prosocial self-sacrifices, which are considered supererogatory behaviours - acts that go beyond our moral duties and cannot be justified recurring to them - (e.g., Slote, 2007). According to their view, to explain this type of behaviours, we must assume that there is a pre-existing moral relationship between people and that responding to the plight of the other is based on automatic, impulsive, or instinctive feelings of responsibility or natural care. Or, as Noddings (2003) put it, when we naturally care about others, we just care about them, no rational effort is required and "want" and "must" become the same thing.

The studies showing that fused people are willing to engage in supererogatory acts for the target of fusion and highlighting the role played by physical, psychological, and emotional activation suggest that strongly fused individuals may be motivated by feelings of responsibility akin to those described by care theorists (e.g., Noddings; Tronto, 1993, 1998); and other additional studies also point to this. For example, identity fusion with extended groups - groups in which people do not personally know each other - occurs when the relational

ties that are established within the context of the family, which are commonly taken as prototypical instances of caring relationships, are projected to the group (Atran et al., 2014; Swann, Buhrmester, et al., 2014; Vázquez et al., 2015; Whitehouse et al., 2014). These ties mediate the relationship between fusion and endorsement of self-sacrifice (Swann, Buhrmester et al., 2014); and degrading them diminishes the level of fusion (Gómez et al., 2019). People fused with a group also perceive themselves as more willing to engage in extreme acts for the group than other group members (Swann et al., 2009) and deem the group as relatively invulnerable, probably because the family-like ties lying at the core of fusion lead them to believe that other group members are also disposed to care and protect the group (Gómez, Brooks et al., 2011). Nonetheless, excluding the research by Swann, Gómez et al. (2014), in which a preliminary approach to the issue using qualitative discourse analysis was made, no study has explored the postulates of care ethics within the fusion context. With this research we aim to address this gap by exploring the relationship between identity fusion, feelings of *visceral responsibility*, and willingness to engage in extreme or supererogatory prosocial behaviour. Following the lead of care theorists, we understand visceral responsibility as an *unconditional, instinctive, and impulsive drive to care, protect, and promote the interests and well-being of the target person or group.*

Identity Fusion, Self-sacrifice, and Shared Feelings of Visceral Responsibility

Strongly fused individuals might be driven to act on behalf of the target of fusion by feelings of visceral responsibility, or by a mental state in which satisfying the needs of the target becomes a goal that motivates behaviour directly, instinctively, and impulsively. This experience of visceral responsibility can explain why strongly fused are so extraordinarily willing to engage in extreme self-sacrifices for the target of fusion; although it can also give rise to the impression that they are invariably impelled to put the interest of the target ahead of their own interest to the highest possible extent, and thus ideally suited for being exploited by

non-caring groups (e.g., groups that may use the benefits of the care that is given to them to harm or subjugate its members). We believe, however, that this impression is not accurate.

Fused individuals are not disconnected from their social environment, and they might use some of the clues provided by it to regulate their feelings of visceral responsibility and their willingness to self-sacrifice, particularly, the clues conveying information about the extent to which the target of fusion also cares. Two different areas of inquiry lead us to assume this.

First, several authors have asserted that family-like relationships, like those established by fused individuals, are characterized by a shared sense of responsibility for one's another welfare (e.g., Earp et al., 2021; Gómez, Brooks et al., 2011). In these relationships, people typically record each other needs and offer non-contingent support to promote the others' welfare (e.g., Clark, et al., 1989, 1998); but they also track the responsiveness of the others to their own needs, experience hurt feelings when the support that they need is not provided to them (e.g., Lemay et al., 2010), and may be even willing to engage in corrective violence under these circumstances (e.g., Fiske & Rai, 2015). On a similar vein, care ethicists have posited that trust and solidarity are needed to care, either because they exist beforehand and allow care (Van Nistelrooij, 2015) or because they result from care-receiving and enable that the caring process keeps going (Tronto, 2013).

Information about the degree in which visceral responsibility is shared with the target of fusion provides crucial cues to ascertain the quality of the relational ties established with it; and to determine whether family-like bonds, trust, and solidarity are being treasured or ignored. Following this, it is reasonable to assume that this information has a regulatory function among strongly fused and that their feelings of visceral responsibility and willingness to self-sacrifice augment when visceral responsibility is shared and diminish when it is not.

Second, descriptive normative information is one of the most powerful sources of social influence (e.g., Berkowitz, 1972; Cialdini & Goldstein, 2004; Deutch & Gerald, 1995). The

cumulative findings of this area of research show that descriptive norms about how similar people behave in a certain situation tend to induce conformity in most individuals (e.g., Cialdini & Trost, 1998). This effect has been widely used to influence prosocial behaviours (e.g., Agerström et al., 2016; Nook et al., 2016; Schultz et al., 2007), and it seems to take place even when people know that the norms are fully arbitrary (Pryor et al. 2019). Moreover, individuals regulate and experience their emotions in ways consistent with their group's emotion norm (Leonard et al., 2011; Weisbuch & Ambady, 2008). And the descriptive norm effect is strong enough to influence general goals in addition to single behaviours and to cross psychological domains, such that observing others' prosocial behaviours induces prosocial emotions and observing others' prosocial emotions induces prosocial behaviours (Nook et al., 2016).

Therefore, visceral responsibility norms may influence strongly fused individuals through conformity processes too.

The idea that strongly fused individuals regulate their responses attending to the extent in which feelings of responsibility and caring are shared has not been formally tested before, but several findings indicate that they may be particularly sensitive to different signs of shared responsibility. In this respect, it has been discovered that perceiving generosity and benevolence as important principles for the group, sharing intense experiences with other group members, and admiring other group members for their personal sacrifices for the group cause identity fusion (Carnes & Lickel, 2018; Gómez et al., 2020; Whitehouse et al., 2017). Also, strongly fused individuals trust more in the target of fusion and are more willing to let the target of fusion sacrifice for them than the weakly fused (Heger & Gaertner, 2018; Vázquez, Gómez, Ordoñana et al., 2017; Whitehouse & Fitzgerald, 2020). The second goal of our research is providing original empirical evidence for the *shared responsibility effect*.

Overview of the Studies

The central question we address here is whether strongly fused individuals are characterized by experiencing strong feelings of visceral responsibility towards the target of fusion, and if such feelings motivate them to self-sacrifice for it. In addition to that, we aim to determine whether descriptive information about the feelings of visceral responsibility experienced by the fusion target influences strongly fused individuals' willingness to self-sacrifice in a direction consistent with it.

To that end, we performed a series of three online studies. We expected that, independently of the target of fusion (an individual, Study 1a, or a group, Study 1b), identity fusion will be positively related to feelings of visceral responsibility and willingness to fight and die for such target. Also, we anticipated that the effect of fusion on willingness to fight and die will be mediated by feelings of visceral responsibility.

And additional study was conducted to offer causal evidence for this underlying process and examine the impact of shared feelings of visceral responsibility. We expected that learning that other ingroup members display low levels of visceral responsibility will attenuate the effects of fusion on willingness to fight and die as compared to learning that other ingroup members display high levels of visceral responsibility.

We did not determine sample size a priori. All the studies were open for a week and then were closed. Unless otherwise stated, participants were recruited using a snowball procedure wherein students from an open university invited their acquaintances to volunteer. The studies were conducted online, via Qualtrics. The materials of the studies and the data that support the findings are publicly available at "Open Science Framework" at https://osf.io/5amvq/https://osf.io/5amvq/

Responses to the measures used in all studies ranged from 1 (*completely disagree*) to 7 (*completely agree*). Participants were asked for the sociodemographic variables at the end.

Studies 1a and 1b

Studies 1a and 1b were designed to test whether individuals strongly fused with their romantic partner (Study 1a), or with their country (Study 1b), display strong feelings of visceral responsibility and willingness to fight and die for them. We expected that identity fusion would be positively related to visceral responsibility and willingness to fight and die. In addition, we anticipated that the positive relation between fusion and willingness to fight and die would be mediated by feelings of visceral responsibility.

As stated by identity fusion theory (Swann et al., 2012), fusion can be extended (with big groups in which people do not personally know most ingroup members, as the country) or local (with small groups as a group of friends or the family). More recently, the theory has been also applied to the ties that people establish with other individuals, values, and several types of entities (Gómez, Chinchilla et al., 2020). We decided to use the country as fusion target because it is an example of fusion with an extended group and the group most used in fusion research. We chose to use the romantic partner as target because it is an example of fusion with one of the most important individuals in our lives.

Method

Participants

Two hundred individuals who were involved in a romantic relationship at the time of the study volunteered to participate in Study 1a (96% Spaniards, 67% women, M_{age} = 36.08, *SD* = 12.81). In Study 1b, the sample was comprised by 182 volunteers (95.6% Spaniards, 67% women, M_{age} = 36.56, *SD* = 13.43).

Sensitivity power analyses. We conducted post-hoc sensitivity power analyses with the online tool developed by Schoemann et al. (2017) to determine the statistical power reached by both studies. Considering our sample sizes and the size of the effects, we found that Study 1a had 100% power and Study 1b had 99% power (see Supplementary Materials).

Procedure

First, *identity fusion* was rated by a reduced and adapted version of the verbal scale of fusion by Gómez, Brooks et al. (2011) including three items ("My partner/country and me are one", "I feel immersed in my partner/country", and "I feel a strong emotional bond with my partner/country", αs = .83 and .86, for studies 1a and 1b respectively). Next, *feelings of visceral responsibility* were measured by a five-item scale developed for the studies ("I feel impulsively obliged to promote the wellbeing of my partner/country", "I feel an unreflective duty to take care of my partner/country", "I feel a duty, based on my deepest feelings, to further the interests of my partner/country", "I feel an unreflective duty to try to stop something bad from happening to my partner/country", and "I feel almost irrationally obliged to act on my partner/country's wishes", αs = .90 and .91, for Studies 1a and 1b respectively). After that, *willingness to fight and die* was evaluated by a reduced and adapted version of the scale by Swann et al. (2009) including three items ("I would fight someone who physically threatened my partner/country", "Hurting other people is acceptable if it means protecting my partner/country", and "I would sacrifice my life if it saved the life of my partner/another member of my country", αs = .75 and .70, for Studies 1a and 1b respectively). In Study 1a, participants were asked too for the number of months they had been in the relationship with his/her partner.

A principal axis factor analysis on the items of identity fusion, feelings of visceral responsibility, and willingness to fight and die with oblique rotation (direct oblimin) showed that each item loaded onto the expected factor, with factor loadings ranging from .629 to .927, in Study 1a, and from .613 to .939, in Study 1b. Additional confirmatory factor analyses revealed that the hypothesized three-factor model fitted the data adequately and has a better adjustment than alternative two-factor and single-factor models in the two studies (see Supplementary Materials).

Results

Table 8 shows the means, standard deviations, and bivariate correlations between the variables. In both studies, the correlations between fusion, feelings of visceral responsibility, and willingness to fight and die were significant, positive, and moderate (between .35 and .44, *p*s < .001). In Study 1a, the number of months in the relationship with the partner did not correlate significantly with any of the other variables, and the analyses controlling for this variable did not alter the results. Consequently, we report the results of the analyses without controlling for time.

To test our hypotheses that the positive relation between identity fusion and willingness to fight and die would be mediated by feelings of visceral responsibility, we performed bootstrapped (*n* boots = 5,000) mediation analysis with PROCESS (Hayes, 2018; Model 4). Identity fusion (mean centred) was included as the predictor, feelings of visceral responsibility as the mediator, and willingness to fight and die as the outcome variable (see Figure 9).

In Study 1a, results yielded significant effects of fusion on visceral responsibility, B = 0.34, $t(198) = 5.33$, p <.001, 95% CI [0.212, 0.461], and willingness to fight and die, B = 0.41, $t(198) = 6.70$, p <.001, 95% CI [0.291, 0.534], such that visceral responsibility and willingness to fight and die for the partner increased as fusion went stronger. The indirect effect of identity fusion on willingness to fight and die for the partner via feelings of visceral responsibility was significant too, B = 0.11, 95% CI [0.054, 0.185].

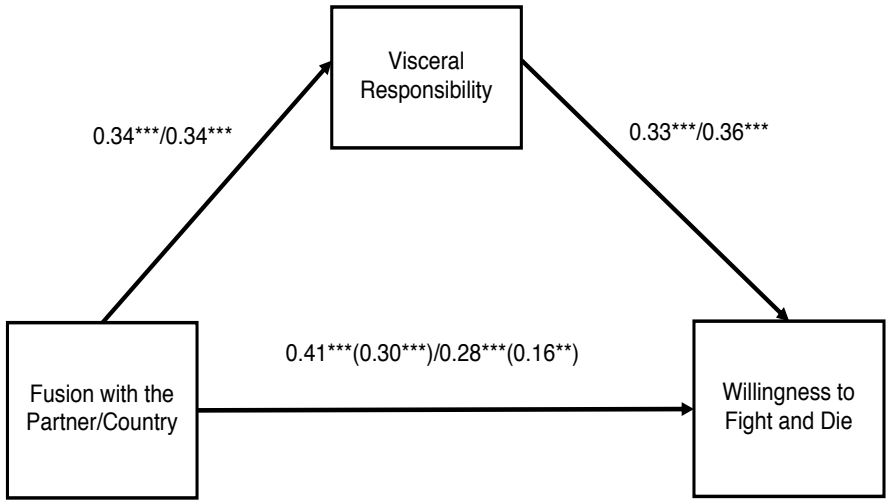
In Study 1b, results showed significant effects of fusion on visceral responsibility, B = 0.34, $t(180) = 6.25$, p <.001, 95% CI [0.233, 0.448], and willingness to fight and die, B = 0.28, $t(180) = 5.02$, $p < .001$, 95% CI [0.172, 0.396], meaning that visceral responsibility and willingness to fight and die for the country increased as fusion went stronger. The indirect effect of identity fusion on willingness to fight and die for the country via feelings of visceral responsibility was also significant, B = 0.12, 95% CI [0.056, 0.207].

Table 8. Descriptive statistics and bivariate correlations between measures (Studies 1 and 2)

Study	Variable	*M*	*SD*	1.	2.	3.
Study 1a	1. Fusion with the Partner	4.28	1.69	-	-	-
	2. Visceral Responsibility	3.11	1.61	.35**	-	-
	3. Fight and Die	4.20	1.63	.43**	.44**	-
	4. Months in the Relationship	117.59	128.53	.03	-.01	.12
Study 1b	1. Fusion with the Country	4.30	1.57	-	-	-
	2. Visceral Responsibility	2.18	1.27	.42**	-	-
	3. Fight and Die	2.23	1.27	.35**	.44**	-
Study 2	A. Low visceral responsibility condition					
	1. Fusion with the Country	3.55	1.45	-	-	-
	2. Visceral Responsibility	2.55	1.30	.41**	-	-
	3. Fight and Die	2.03	1.15	.25**	.51**	-
	B. High visceral responsibility condition					
	1.Fusion with the Country	3.71	1.66	-	-	-
	2. Visceral Responsibility	2.85	1.58	.71**	-	-
	3. Fight and Die	2.38	1.34	.51**	.63**	-
	C. Correlations across experimental conditions					
	1. Fusion with the Country	3.63	1.55	-	-	-
	2. Visceral Responsibility	2.69	1.45	.57**	-	-
	5. Fight and Die	2.20	1.25	.39**	.58**	-

** p <.01

Figure 9. Indirect effects of identity fusion on willingness to fight and die for the partner/country via feelings of visceral responsibility (Studies 1a and 1b)



Discussion

As expected, the results indicated that identity fusion was positively related to feelings of visceral responsibility and willingness to fight and die. Also as predicted, the effect of fusion on fight and die appeared to be mediated by feelings of visceral responsibility. The results were replicated independently of the target of fusion, an individual (Study 1a, the partner), or a group (Study 1b, the country).

These results are consistent with our thesis that one of the mechanisms through which identity fusion motivates extreme self-sacrifice is via feelings of visceral responsibility. However, although promising, the previous studies concur on the same limitation, as it is that the correlational nature of the design does not allow to establish causality. Study 2 was conducted to solve this limitation by manipulating feelings of visceral responsibility.

Study 2

As recommended by Spencer et al. (2005), we complement our findings of correlational mediation in Studies 1a and 1b by manipulating the hypothesized mediator through bogus feedback about the extent in which ingroup members feel viscerally responsible towards the group. For this purpose, after assessing identity fusion with the country, we told participants either that most ingroup members feel visceral responsibility towards the country or that most ingroup members do not feel visceral responsibility towards the country.

We hypothesized that, as compared to telling participants that most ingroup members feel visceral responsibility towards the country, telling participants that most ingroup members do not feel visceral responsibility towards the country would attenuate feelings of visceral responsibility and willingness to fight and die among strongly fused participants. In accordance with previous findings showing that the devotion to the group of fused individuals goes beyond the devotion manifested by their fellow group members (e.g., Swann et al., 2009), we anticipated positive relationships between fusion and feelings of visceral responsibility, and between fusion and willingness to fight and die, across conditions.

Method

Participants

Three hundred three Spaniards volunteered for the study (60.4 % women*;* M_{age} = 35.08; *SD* = 13.35).

Sensitivity power analysis. We conducted a post-hoc sensitivity power analysis with G*Power (Faul et al., 2007) for lineal multiple regression (fixed model, R^2 deviation from zero) to determine the statistical power reached by Study 2 considering the effect size found in the moderated multiple regression on willingness to fight and die (f^2 = .23). With our sample size and three predictors (identity fusion, experimental condition, and the two-way interaction), Study 2 had 100% power.

Procedure

First, participants responded to the same measure of *fusion with the country* used in previous studies (α = .84). After that, all participants read the following text: "In some previous studies conducted by our research team, we have presented thousands of Spaniards a series of questions to assess their feelings of visceral responsibility towards Spain. People who feel viscerally responsible for their country are willing to protect and defend it because their deepest feelings and emotions impel them to do so; and we wanted to know whether the Spaniards feel this type of visceral responsibility or not."

Participants were then randomly assigned to one of the two experimental conditions. In the *high visceral responsibility condition (n* = 145), participants learned that most Spaniards feel viscerally responsible towards Spain, and that the feelings that they experience is what impulse them in an almost irrational way to do things to protect and defend Spain. In the *low visceral responsibility condition (n* = 158), participants learned that most Spaniards do not feel viscerally responsible towards Spain, and that the feelings that they experience do not impulse them to do things to protect and defend Spain.

As manipulation check, participants were asked to complete the measure of *feelings of visceral responsibility* (α = .92). Finally, they were requested to answer to the *willingness to fight and die* scale (α = .74) from Studies 1a and 1b.

A principal axis factor analysis on the items of identity fusion, feelings of visceral responsibility, and willingness to fight and die with oblique rotation (direct oblimin) revealed the expected three factors and showed that each item loaded onto the corresponding factor, with factor loadings ranging from .420 to 1.02. Additional confirmatory factor analyses revealed that the hypothesized three-factor model fitted the data adequately and has a better adjustment than alternative two-factor and single-factor models (see Supplementary Materials).

Results

Table 8 shows the bivariate correlations and descriptive statistics within the low visceral responsibility (Panel A), the high visceral responsibility (Panel B), as well as across conditions (Panel C). All the variables were positively related.

A preliminary t-test showed that, as anticipated, the level of fusion with the country did not vary as a function of the experimental condition, $t(301) = -0.88$, $p = .378$, 95% CI [-0.508, 0.193].

Moderation Analyses

To determine whether our manipulation worked as expected, we performed bootstrapped (*n* boots = 5,000) moderated multiple regressions using the Macro PROCESS (Hayes, 2018; Model 1), including identity fusion (mean centered) as predictor, experimental condition (0 = low visceral responsibility; 1 = high visceral responsibility) as moderator, and feelings of visceral responsibility as outcome variable. The same analyses were also performed to test the effects on the outcome variable, but regressing willingness to fight and die on the predictor variables.

Feelings of visceral responsibility. Results yielded a significant interaction between identity fusion and experimental condition, $B = 0.30$, $t(299) = 3.51$, $p < .001$, 95% CI [0.133, 0.473], meaning that the effect of identity fusion on feelings of visceral responsibility towards the country varied as a function of experimental condition. Simple slope analysis revealed that identity fusion was positively related to feelings of visceral responsibility in both the high visceral responsibility condition, $B = 0.67$, $t(299) = 11.55$, $p < .001$, 95% CI [0.559, 0.788], and the low visceral responsibility condition, $B = 0.37$, $t(299) = 5.79$, $p < .001$, 95% CI [0.244, 0.496], but the relationship between both variables was weaker in the condition of low visceral responsibility (for details see Table 9 and Figure 10).

Additional analyses showed as well that the effect of experimental condition was significant among strongly fused participants, $B = 0.73$, $t(299) = 3.68$, $p < .001$, 95% CI [0.341, 1.125], who feel more viscerally responsible for the country in the condition of high visceral responsibility. However, the effect of condition was not significant among weakly fused participants, $B = -0.28$, $t(299) = -1.44$, $p = .152$, 95% CI [-0.660, 0.103].

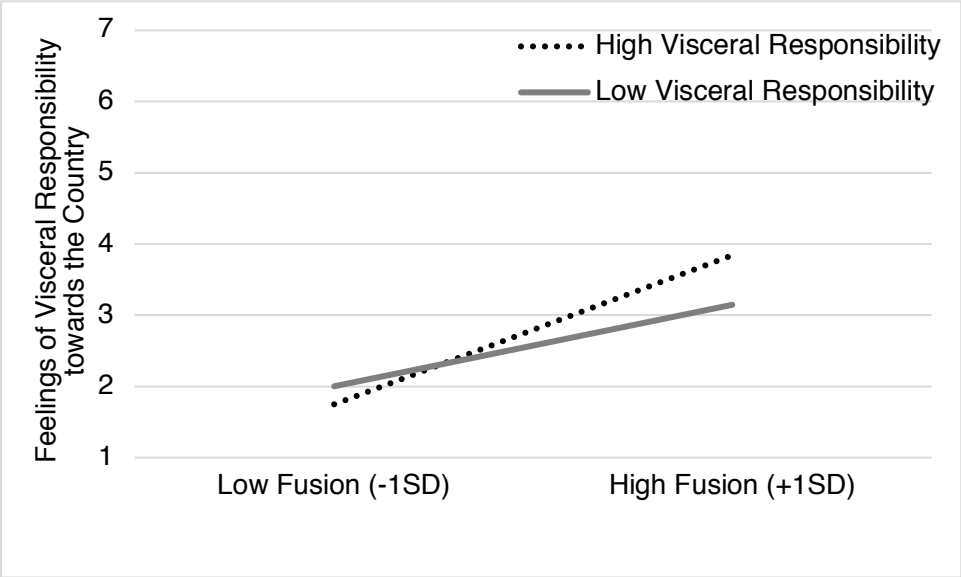
The main effect of identity fusion was significant too, $B = 0.37$, $t(299) = 5.79$, $p < .001$, 95% CI [0.244, 0.496], indicating that feelings of visceral responsibility increased as fusion went stronger. The effect of experimental condition was not significant, $B = 0.21$, $t(299) = 1.61$, $p = .107$, 95% CI [-0.047, 0.478].

Table 9. Moderated regression on feelings of visceral responsibility (Study 2)

Predictor	*B*	*SE*	*t*(299)	*p*	LLCI	ULCI
Constant	2.58	0.09	27.91	.000	2.395	2.759
Fusion	0.37	0.06	5.79	.000	0.244	0.496
Condition	0.21	0.13	1.61	.107	-0.047	0.478
Fusion X Condition	0.30	0.09	3.51	.000	0.133	0.473
Simple slopes:						
High Visceral Responsibility	0.67	0.05	11.55	.000	0.559	0.788
Low Visceral Responsibility	0.37	0.06	5.79	.000	0.244	0.496
High Fusion (+1 SD)	0.73	0.20	3.68	.000	0.341	1.125
Low Fusion (-1SD)	-0.28	0.19	-1.44	.152	-0.660	0.103

Note: $R^2 = .36$

Figure 10. Feelings of visceral responsibility as a function of identity fusion and experimental condition (Study 2)



Willingness to fight and die. Results showed a significant interaction between identity fusion and experimental condition, $B = 0.21$, $t(299) = 2.53$, $p = .012$, 95% CI [0.047, 0.380], indicating that the effect of identity fusion on willingness to fight and die for the country varied as a function of experimental condition. Simple slope analysis showed that fusion was positively related to willingness to fight and die in both the high visceral responsibility condition, $B = 0.41$, $t(299) = 7.19$, $p < .001$, 95% CI [0.297, 0.521], and the low visceral responsibility condition, $B = 0.19$, $t(299) = 3.14$, $p = .002$, 95% CI [0.073, 0.318], although the relationship between both variables was smaller in the condition of low visceral responsibility (for more details see Table 10 and Figure 11).

Additional analyses revealed too that the effect of experimental condition was significant among strongly fused participants, $B = 0.67$, $t(299) = 3.42$, $p = .001$, 95% CI [0.283, 1.049], who were more willing to fight and die for the country in the condition of high visceral responsibility. Contrarily to that, the effect of condition was not significant among weakly fused participants, $B = -0.05$, $t(299) = -0.24$, $p = .806$ 95% CI [-0.419, 0.326].

The main effect of identity fusion was also significant, $B = 0.19$, $t(299) = 3.14$, $p = .002$, 95% CI [0.073, 0.318], meaning that willingness to fight and die for the country increased as fusion went stronger. Lastly, the main effect of experimental condition was significant too, $B = 0.30$, $t(299) = 2.31$, $p = .021$, 95% CI [0.045, 0.558], revealing that willingness to fight and die was stronger in the condition of high visceral responsibility than in the condition of low visceral responsibility.

Table 10. Moderated regression on willingness to fight and die (Study 2)

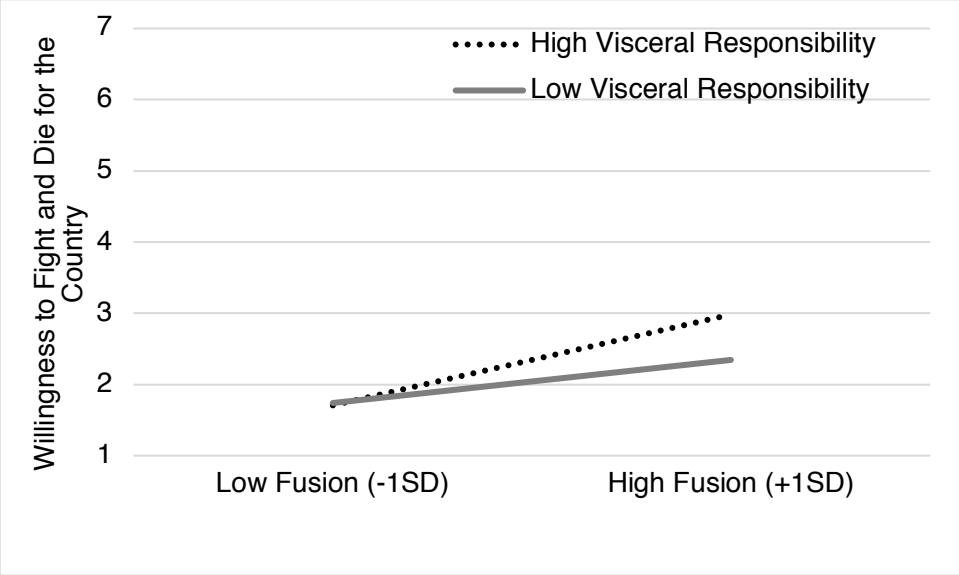
Predictor	*B*	*SE*	*t*(299)	*p*	LLCI	ULCI
Constant	2.05	0.09	22.70	.000	1.869	2.234
Fusion	0.19	0.06	3.14	.002	0.073	0.318
Condition	0.30	0.13	2.31	.021	0.045	0.558
Fusion X Condition	0.21	0.08	2.53	.012	0.047	0.380
Simple slopes:						
High Visceral Responsibility	0.41	0.06	7.19	.000	0.297	0.521
Low Visceral Responsibility	0.19	0.06	3.14	.002	0.073	0.318
High Fusion (+1SD)	0.67	0.19	3.42	.001	0.283	1.059
Low Fusion (-1SD)	-0.05	0.19	-0.24	.806	-0.419	0.326

Note: $R^2 = .19$

Mediational Analysis. Lench et al. (2014) advice conducting mediational analysis with manipulation checks to correctly determine the causal role of the intended mental state in generating the outcome. Following their suggestion, we performed bootstrapped (boots = 5,000) mediational analysis with the Macro PROCESS (Hayes, 2018; Model 8), including identity fusion (mean centred) as predictor, experimental condition (0 = low visceral responsibility; 1 = high visceral responsibility) as moderator, feelings of visceral responsibility (mean centred) as mediator, and willingness to fight and die as outcome variable (see Figure 12). Results confirmed that the overall indirect effect of the interaction between identity fusion

and experimental condition on willingness to fight and die through feelings of visceral responsibility was significant, $B = 0.13$, 95% CI [0.048, 0.229]. Additionally, the indirect effect of identity fusion on willingness to fight and die for the country via visceral responsibility was significant in both the high visceral responsibility condition, $B = 0.30$, 95% CI [0.207, 0.394], and the low visceral responsibility condition, $B = 0.16$, 95% CI [0.092, 0.246].

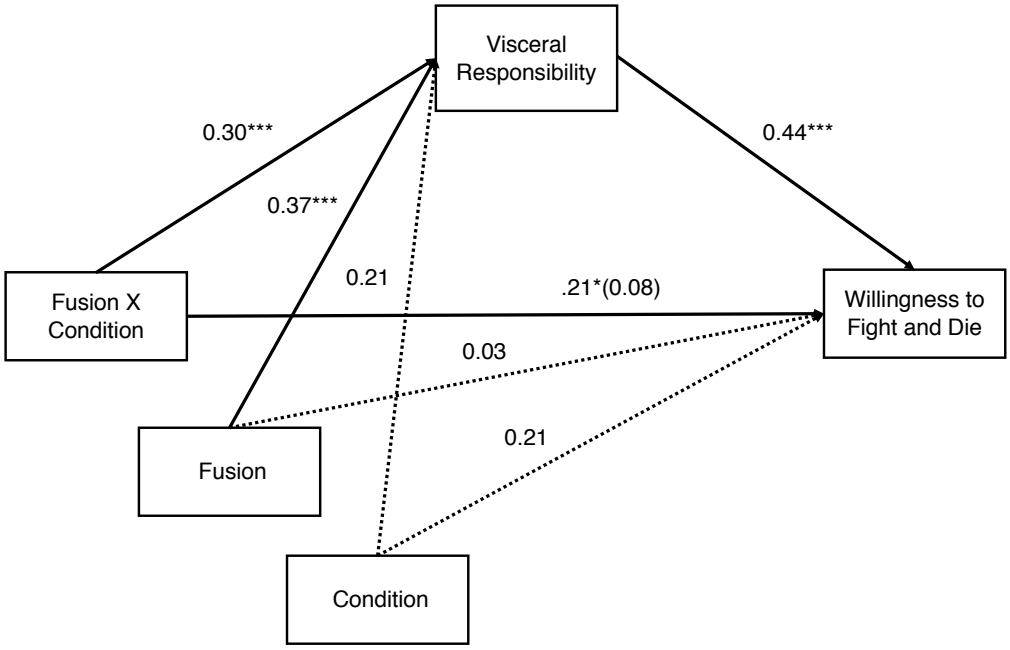
Figure 11. Willingness to fight and die as a function of identity fusion and experimental condition (Study 2)



Discussion

The results of Study 2 showed that manipulating visceral responsibility influenced willingness to fight and die for strongly fused individuals. Those who were informed that feelings of visceral responsibility are low among ingroup members reported less willingness to fight and die for the group than those who were informed that feelings of visceral responsibility are high.

Figure 12. Indirect effects of the interaction between identity fusion and experimental condition on willingness to fight and die for the country via feelings of visceral responsibility (Study 2)



***$p < .001$, * $p < .05$

General Discussion

The present research examined, for the first time, the relationship between identity fusion, feelings of visceral responsibility, and extreme prosocial behaviour. In two studies considering fusion with an individual (the romantic partner, Study 1a) or with a group (the country, Study 1b), we found that strongly fused participants experienced more feelings of visceral responsibility towards the target of fusion and were more willing to fight and die for it; and that the relationship between fusion and willingness to fight and die seemed to be mediated by visceral responsibility. After that, we manipulated participants' feelings of visceral responsibility towards the country by providing them with bogus information about the feelings of visceral responsibility experienced by their fellow citizens (Study 2); and found

that strongly fused participants who learned that visceral responsibility is not felt by most ingroup members were less willing to fight and die for the country as compared to those who learned that feelings of visceral responsibility are widely shared. Moreover, the relationship between identity fusion and willingness to fight and die was mediated by feelings of visceral responsibility in the two experimental conditions.

Our results make contributions to different domains. Theoretically, our studies advance identity fusion research and tie it to care ethics in at least three different ways. First, by showing that strongly fused individuals are willing to fight and die because they feel viscerally responsible to protect and promote the wellbeing of the target of fusion, they highlight a new mediational mechanism not examined before: feelings of visceral responsibility. Our studies complement previous research on the role of physical, psychological, and emotional activation as explicative mechanisms of the effect of identity fusion on extreme pro-group behaviours (Gómez, Brooks et al., 2011; Swann et al., 2009; Swann, Gómez et al., 2014; Swann, Gómez, Huici et al., 2010). The findings manifest that, without need for further abstract justification, fusion may generate automatic, impulsive, and instinctive feelings of visceral responsibility that motivate strongly fused to care and defend the interests of the target of fusion at a high cost to the self, as anticipated by some descriptions of responsibility or natural care (e.g., Noddings, 2003; Tronto, 1993)[4]. Second, our studies add to previous research on the moderators that weaken fusion effects on extreme self-sacrifices, like knowing that self-sacrifices can be reasonably delayed or holding doubts about the own degree of fusion (Paredes et al., 2018, 2019), revealing that strongly fused individuals' willingness to engage in extreme pro-group behaviours diminishes when they learn that other group members are not committed

[4] It is also the first time that a mediational mechanism of the effects of fusion has been explored out of the domain of fusion with a group, and the first case where a mediator has been demonstrated to work with fusion with different types of entities - i.e., individuals and social groups -.

to care and protect the group's interests. These results are consistent with the discovery that degrading either relational ties - sentiments towards individual group members - or collective ties – sentiments towards the group as a whole - diminishes identity fusion (Gómez et al., 2019); and they lend indirect support to the idea that strongly fused individuals perceive themselves as essentially imbued in caring relationships and that some shared responsibility or solidarity on the part of the target of fusion is needed to sustain a high level of care, as envisioned by care theorists too (e.g., Noddings, 2003; Tronto, 2013; van Nistelrooij, 2015). Third, and also following the assumptions of care ethics (e.g., Slote, 2007; van Nistelrooij, 2015), by showing that although to a lesser extent, identity fusion predicts willingness to fight and die even when participants are told that most ingroup members do not feel viscerally responsible for the group, our studies indicate that the extreme pro-group behaviours typical of the strongly fused might not rest entirely on strict reciprocity and not disappear when the expectation of mutualistic compensation breaks down. It seems then that the moral machinery sustaining extreme prosocial behaviour among the strongly fused is fuelled by the passions and social instincts that regulate caring relationships rather than by abstract moral principles or imperatives, which might not have the strength needed to motivate costly personal sacrifices (e.g., Slote, 2007).

At a practical level, our studies have several implications as well. Because people generally fuse with groups that do not comprise everyone, identity fusion entails a strong risk of moral parochialism, which can be particularly dangerous in the case of individuals fused with groups that legitimize the use of violence, such as terrorists (e.g., Chinchilla et al., 2021, Gómez et al., 2017, 2020). Relating to that, our research points to several paths on which program developers and practitioners could embark to prevent the problems related to parochialism. For instance, when dealing with individuals who are already fused with violent groups, the deployment of persuasive messages to challenge the assumption that feelings of

visceral responsibility are shared withing the group might lessen their proneness to engage in violence. Although the use of this strategy alone might not suffice to deter aggression, it has the advantage that it can be easily combined with other actions without substantially increasing the costs of interventions. In general, it might also be useful to develop programs aimed at promoting fusion with inclusive groups characterized by highly permeable and flexible borders. In this respect, Talaifar and Swann (2018) have posited that we should work to foster fusion with humanity because that can give us less reasons to fight one another and more reasons to work together towards mutually beneficial outcomes.

Undoubtedly, our research has limitations as well. Our studies do not provide direct evidence in support of the causal role of identity fusion on feelings of visceral responsibility. It is much more plausible that the feelings of oneness that the fused experience take precedence and cause visceral responsibility than the other way around, but future research should address this issue by collecting longitudinal data and/or experimentally manipulating fusion.

Our studies signal some future lines of research that could be pursued too. One interesting option would be testing whether the mediational mechanism that we have discovered generalizes to mindless entities, like values, objects, or firms. The fact that previous research has extended the scope of identity fusion from groups to other targets as individuals, values, etc., does not mean that the underlying mechanisms that increase, decrease, or mediate the effects of fusion are identical in all cases. While we have demonstrated that visceral responsibility mediates the effect of fusion on self-sacrifices considering fusion with a group and with an individual, it is not entirely clear whether the same mechanism also operates when fusion is felt towards entities that do not have mind. We suspect that fused individuals also feel visceral responsibility towards such entities but that these feelings are not based on the same processes that are important in the case of fusion with sentient beings – e.g., the expectation of causing or avoiding suffering or pain - unless they anthropomorphize them (e.g., Epley et al.,

2007). Also, other studies could explore the specific emotions that give rise to feelings of visceral responsibility among the fused. We think that the most obvious candidate is empathy or sympathetic concern (e.g., Batson, 2019; Slote, 2007). Lastly, future research should explore as well if the findings of our studies apply to other types of relational bonds, such as those based on communal sharing (Fiske, 1992).

Conclusions

The instinct for self-preservation is one of the most powerful drivers of behaviour, but it can be overpassed by our want to help. From abnegate comrades and loyal friends to old and current terrorists and social activists, countless of examples illustrate that humans are often willing to incur in extreme personal sacrifices, and even to die, for the people and groups cherished to them. Such examples exceed our moral expectations and are so perplexing that we have been wondering what kind of psychological processes might explain this type of acts during millennia. The path to a definitive answer to this question is still long; but the evidence accumulated let us make some empirically based conjectures about what we would find by digging into the mind of individuals who engage in extreme self-sacrifices on others' behalf. As the quote that opens this paper illustrates, we will probably discover that when our reality merges with that of the other and we start to feel its reality, our passions also compel us to act on behalf of the other as if it was in our own behalf, giving rise to feelings of visceral responsibility that motivate us to engage in significant personal sacrifices. In the context of social psychology, this feeling of utmost closeness has been captured in the most faithful way by the construct of identity fusion; and the moral machinery lying at its core has been described by the ethics of care.

CHAPTER 6

General Discussion

This thesis had to main goals. First, we wanted to write a short narrative review article including an overview of the theoretical and empirical contributions to identity fusion theory since 2015; addressing the principal doubts, misconceptions, and untested assumptions about the construct; and signalling some pertinent future lines of research. Second, we wanted to blend identity fusion theory with some of the findings of moral psychology to reach a deeper knowledge of the factors that cause identity fusion, moderate its consequences, and mediate its effects. Specifically, we wanted to determine whether (1) admiration leads to identity fusion; (2) beliefs about the moral justifiability of violence regulate the type of behaviour (violent vs. nonviolent) in which strongly fused individuals decide to engage; and (3) strongly fused individuals engage in extreme prosocial behaviours because they feel visceral responsibility towards the fusion target.

In what follows, we outline the results of our investigation, address its main theoretical and practical contributions, signal its principal limitations, and point out novel hypotheses and directions for upcoming research.

Results and Contributions

Narrative Review Article

The narrative review included in Chapter 2 covers the theoretical and empirical studies conducted from 2015 to 2020 and reveals that, during this period, research on identity fusion has been quite prolific. Multidisciplinary teams of scientists from the five continents conducted numerous laboratory and field studies to extend the scope of the theory and apply it to special populations of interest, like hooligans, political partisans, martial arts practitioners, college fraternity/sorority members, military veterans, or terrorists, among others (e.g., Gómez et al., 2017; Kapitány et al., 2019; Newson et al., 2018; Whitehouse et al., 2017). Their efforts led to notable findings about the nature, causes, underlying mechanisms, and applications of identity fusion. In fact, this is one of the rare cases withing the psychosocial sciences wherein a theory

has been created, expanded multidisciplinarily and transnationally, and put to the test with an ample variety of samples of unquestionable socio-political interest in ecologically valid settings in such a brief amount of time - just around 10 years. Identity fusion has been also recently incorporated into two major theories to explain radicalization leading to extremism: the devoted actor model and the 3N model of radicalization. The *devoted actor model* applies to individuals fused with a group with which they share sacred values - values non fungible or negotiable - (Gómez et al., 2017; Sheikh et al., 2016; Vázquez et al., 2020). The *3N model* integrates identity fusion together with needs - i.e., significance - and violence-justifying narratives (e.g., Kruglanski, Bélanger et al., 2019).

Regarding the *nature of identity fusion,* one of the most notable advances was the extension or application of the feelings of fusion to other targets, such as other individuals (Joo & Park, 2017; Kunst et al., 2019; Vázquez, Gómez, Ordoñana et al., 2017; Walsh & Neff, 2018), animals (Buhrmester et al., 2018), trademarks (Hawkins, 2019), or values and convictions (Fredman et al., 2017). Other important finding relates to the fact that school-age children - from 6 to 12 years - are uncapable of experiencing those aspects of identify fusion that depend on the development of a mature sense of the personal identity, although they can feel viscerally connected to a group and express willingness to engage in sacrifices for it; and thus are able to experience a state that has been labelled as "protofusion" (Gaviria et al., 2015). Lastly, research showed too that the relational ties - attachment to individual group members - and the will to engage in extreme behaviours for the group characteristic of strongly fused individuals are more resistant to change after internal events threatening the group than the collective ties - attachment to the group as a whole - (Vázquez et al., 2017).

Concerning the *causes of identity fusion*, research accumulated a remarkable body of evidence showing that sharing essential biological characteristics and intense positive or negative emotional experiences with other group members causes identity fusion (e.g., Atran

et al., 2018; Kavanagh et al., 2018; Segal et al., 2018; Vázquez, Gómez, Ordoñana et al., 2017; Whitehouse et al., 2017). Other studies provided evidence for additional mechanisms as well, indicating, among others, that intergroup conflicts that collide with people's political convictions may promote fusion with groups to which the individual does not belong (Kunst et al., 2019); or that sharing central moral values or convictions with other group members increases fusion (Carnes & Lickel, 2018).

As for the *consequences of fusion*, abundant studies demonstrated that, although the specific type of behaviour that the fused exhibit varies as a function of the target of fusion and the circumstances, identity fusion drives individuals to care, protect, and defend the entities with which they are fused. Additional studies also showed that identity fusion influences moral and socio-political views and behaviour (e.g., Kunst et al., 2019; Talaifar & Swann, 2018), and may increase well-being (e.g., Besta et al., 2018; Walsh & Neff, 2018).

Regarding the *underlying mechanisms associated to identity fusion*, research uncovered several factors that moderate and mediate its effects. For instance, some studies showed that strongly fused individuals are more inclined to engage in extreme pro-group behaviours when they feel morally obliged to do so (Kunst et al., 2018); and other studies demonstrated that feelings of self-expansion and group efficacy beliefs sequentially mediate the effects of fusion (Besta et al., 2018).

Finally, the *applications of identity fusion theory* were mostly focused on the development of techniques for preventing violent extremist offences caused by this visceral feeling of connection. In this respect, Gómez et al. (2019) conducted several studies manifesting that degrading either relational ties or collective ties diminishes identity fusion and willingness to engage in extreme pro-group actions.

The research conducted from 2015 to 2020 has considerably increased our knowledge about identity fusion, but it has been accompanied by a series of misconceptions and untested assumptions that were addressed in the review too.

The main *misconceptions about identity fusion* are related to its nature as well as to its antecedents and consequences. Two of the most important misconceptions about the nature of identity fusion are the assimilation of identity fusion to two related but different constructs: group identification (Tajfel & Turner, 1979) and communal sharing (Fiske, 1992). Because the three constructs refer to the psychological ties that individuals establish with social groups, they are intimately related, but group identification and communal sharing entail disregarding the individual identities of other group members and treating them as equivalent and undifferentiated (e.g., Fiske, 1992; Hogg et al., 1993, 1995; Hogg, & Hardie, 1991), whereas identity fusion implies recognizing their unique personal identities along with their social identities (Swann et al., 2009). Other common misunderstanding about the nature of fusion consists of assimilating the construct to a personal disposition or a personality trait that motivates individuals to fuse with groups rather than understanding it as something that results from intragroup processes and dynamics, as research has shown (e.g., shared experiences, Jong et al., 2015; Whitehouse et al., 2017). On the other side, the principal misconceptions about the antecedents and consequences of fusion consist of assuming that it (1) is the single or more relevant predictor of extreme self-sacrificial behaviour (e.g., Whitehouse, 2018); (2) cannot motivate sacrifices on behalf of groups to which individuals do not belong (Crimston & Hornsey, 2018); or (3) is a fleeting state from which people enter and exit at will to evaluate the personal risks and benefits associated with self-sacrifice (Wiessner, 2018).

Finally, the key *untested assumptions about identity fusion* are the idea that it is uniquely, or even inherently, related to pro-group violence and negative behaviours (e.g., Kipper & Sosis, 2018; Newson, 2018); and that it only relates to reported willingness to die for

the group, but not to real willingness to die for it (Lankford, 2018). In the studies covered in Chapter 4, we demonstrate that the type of behaviours (violent vs. non-violent) in which the strongly fused choose to engage depends on ideological factors, concretely, on their moral convictions.

Our narrative review helps curious readers to gain a deeper knowledge about identity fusion and the causal mechanisms that might motivate individuals to risk their lives and physical integrity for the entities that are cherished to them as well as to discard common misconceptions that have appeared withing the field and, as such, it works like a map that aids interested people to navigate through the field.

Given that identity fusion has become a very flourishing field, several additional findings about the nature of identity fusion and its antecedents, underlying mechanisms, and consequences have accumulated since we published the paper included in Chapter 2. For instance, it has been found, among others, that perceived brand social investment and extrinsic customer investment may engender brand identity fusion (Krishna & Kim, 2021); that obsessive passion can enhance the degree of fusion of entrepreneurs with their organization (Lee et al., 2021); that perceptions of national continuity may foster identity fusion with national groups (Shiromaov et al., 2021); and that moral convergence may engender fusion with radical groups (Atari et al., 2021). It has also been found that identity fusion evokes empathic concern and personal distress when the group is at risk and that empathic concern, in turn, augments strongly fused individuals' proneness to self-sacrifice, whereas feelings of personal distress inhibit it (Landabour et al., 2021). Paredes et al. (2021) found that strongly fused individuals are more willing to self-sacrifice when the situation is more extreme; Martel et al. (2021) discovered that identity fusion is a better predictor of extreme behaviour than moral mandates and sacred values; and Wolfowick et al. (2021) conducted a meta-analysis

including 50 risk factors for violent radicalization and found that identity fusion is the most potent predictor of radical intentions.

Admiration and Identity Fusion

Chapter 3 summarizes the results of six studies exploring the role of feelings of admiration as a causal antecedent of identity fusion and self-sacrificial behaviour for group or values.

Study 1 is a quasi-experimental study conducted within the Spanish prisons. It includes a sample of Jihadist terrorists and a sample of non-Jihadist Muslim inmates as comparison group and reveals that, as compared to non-Jihadist criminals, Jihadist terrorists experience more admiration towards radical Islamist groups which, in turn, is associated to greater fusion with religion and greater willingness to engage in costly sacrifices for religion in prison.

Studies 2-6 are experimental studies that manipulate feelings of admiration and bring causal evidence in support of the mechanism discovered in prison. Consistently with research showing that admiration is an emotion elicited by virtue (e.g., Cox, 2010; Schnall et al., 2010; Vianello et al., 2010; Vyver & Abrams, 2015), they also show that the effect of admiration on identity fusion and self-sacrifice is due to an increase in the cognitive salience of the costly sacrifices that the admired person made for the group. The studies examine the boundary conditions under which admiration exerts its effects as well.

Study 2 shows that thinking about an admired ingroup member sequentially increases the cognitive salience of the sacrifices that this person made for the group and identity fusion with the group which, in turn, augment willingness to fight and die for it. Study 3 extends the previous findings showing that the impact of admiration on identity fusion and willingness to fight and die is not moderated by whether admiration is felt towards a group or an individual. Study 4 reveals that, although thinking of an admired outgroup member also increases the cognitive salience of the sacrifices made for the group, identity fusion with the ingroup, and

willingness to fight and die for it as compared to thinking about non-admired individuals, the effects are smaller than those of thinking about an admired ingroup member. Study 5 shows that thinking about an ingroup member admired for his/her personal sacrifices for the group increases the cognitive salience of the sacrifices, identity fusion, and willingness to fight and die; whereas thinking about an ingroup member admired for his/her personal qualities has weaker or non-significant effects. Lastly, Study 6 shows that thinking about an admired ingroup member impacts actual pro-social behaviours towards other group members and that the effects are still present one month after the experimental manipulation.

The results of these studies make important theoretical contributions to identity fusion theory and research on admiration. First, our results bring strong support to the idea that admiration is an other-praising emotion triggered by the observation of virtuous acts (Algoe & Haidt, 2009; Onu et al., 2016), and illustrate that this emotion produces a concatenation of effects that motivate individuals to make extreme self-sacrifices for their groups or ideological convictions. Admiration produces this effect by augmenting the cognitive salience of the sacrifices that others have made for the group which, in turn, increases identity fusion and ultimately boosts willingness to engage in personal sacrifices. Importantly, our results challenge previous research on admiration by showing that the consequences of this emotion on intergroup relations are not invariably positive, as has been widely assumed (e.g., Onu et al., 2016). Admiration can make individuals more prone to defend and protect their cherished group or values by all means necessary, including trough violence, and thus this emotion can also inspire evil actions. Second, our studies complement previous research on the role of shared intense emotional experiences as causal antecedents of identity fusion (e.g., Jong et al. 2015; Kavanagh et al., 2018; Segal et al., 2018; Whitehouse, 2017) by manifesting that other-directed positive emotions experienced in solitude can also enhance visceral feelings of oneness with groups, or ideological convictions, and promote extreme behaviours on their

behalf. Third, our studies bring to light several factors that moderate the effects of admiration on identity fusion and self-sacrifices, showing that the impact of this emotion is stronger when it is experienced towards ingroup members and when is due to the personal sacrifices that the admired person has made for the group. Fourth, our research demonstrates that admiration has real, long-term behavioural consequences, that go beyond momentary priming in an experimental setting and generalize to group members towards whom this emotion is not directly felt.

Our studies have several practical implications to prevent radicalization leading to violence and promote the deradicalization or disengagement. Regarding the prevention of radicalization, previous research suggests that feeling admiration towards groups that legitimate aggressive behaviours may be a risk factor for engagement in extreme violent acts (e.g., Atran, 2010; Kruglanski, Bélanger et al., 2019). One strategy to counter the influence exerted by such groups could be to ensure that people, especially vulnerable individuals, have access to prosocial groups, inspiring role models, and everyday heroes that blatantly oppose the use of violence and work to promote the well-being and interests of all society members, such as youth centres, charitable organizations, social activists, or volunteers. Concerning the deradicalization and disengagement from violence, our results indicate that undermining feelings of admiration towards terrorist groups can be effective to offset the impulse to emulate the actions of terrorists. One possible strategy to reach this effect might be eroding the moral image projected by violent extremist groups by highlighting their moral hypocrisy - the inconsistences that exist between their moral principles and actions. Giving public exposure to the voices of former extremist offenders who have defected from the group because of such hypocrisy may be an invaluable asset in this respect (e.g., Ashour, 2010; Jacobson, 2009; Speckhard et al., 2018; Tapley & Clubb, 2019).

Beliefs in the Moral Justifiability of Violence and Identity Fusion

Chapter 4 presents two cross-sectional studies, focusing on fusion with a local (the family; Study 1) or an extended group (the country; Study 2), that examine whether beliefs in the moral justifiability of violence moderate the relationship between identity fusion and willingness to engage in extreme violent and nonviolent pro-group behaviours when the group is under a serious threat. Results of both studies showed that strongly fused individuals are more willing to engage in pro-group violence than weakly fused individuals, but only when they believe that violence is morally justifiable. Among those individuals who did not endorse the idea that violence is justifiable, identity fusion does not predict willingness to use violent means. Additionally, the moderating role of beliefs in the moral justifiability of violence does not extend to willingness to engage in extreme nonviolent acts, which only increase as the degree of fusion augments.

These results make several theoretical and practical contributions. At a theoretical level, our studies complement identity fusion in at least two relevant ways. First, they constitute the first systematic exploration of the effects of explicit moral beliefs on fused individuals' proneness to engage in pro-group violence. They show that whether fused individuals engage or not in pro-group violence depends on what they believe about its moral justifiability and, by doing so, they help to disentangle the true nature of identity fusion and solve some doubts about the relationship between fusion and violence that have appeared withing the field (Kiper & Sosis, 2018; Swann & Buhrmester, 2015). Highly fused individuals are strongly motivated to defend and protect the group, but they are not necessarily more inclined to engage in violence than weakly fused individuals and do not always manifest a violent side.

Second, our studies complement previous research on the mechanisms that moderate the behavioural consequences of identity fusion, such as impulsive decision making (e.g., Joo & Park, 2017), knowing that the own personal sacrifices could be reasonable delayed (Paredes et al., 2018), or holding meta-cognitive doubts about the own degree of fusion (Paredes et al.,

2020). Whereas these factors have an indiscriminate impact on pro-group self-sacrifices, our studies identify a novel mechanism that may cancel out fusion effects on violence without influencing the relationship between identity fusion and extreme nonviolent pro-group acts and illustrate the extraordinary weight that moral convictions have in regulating violence directed by impulses and visceral passions. In addition to that, our studies also advance moral psychology and research on the two systems of moral regulation (Janoff-Bulman et al., 2009) highlighting that the proscriptive moral system might be stronger than the prescriptive system even among those individuals for whom the defence and protection of the of the group is experienced as a moral mandate, as seems to be the case among strongly fused.

At a practical level, our studies suggest some strategies that may be used by educators and program developers and practitioners to fight violence motivated by feelings of oneness with a group. In contrast to other measures suggested before, which mostly focus on breaking the ties that bind individuals to the group and its members (e.g., Gómez et al., 2019), this new measure is only focused on the change of moral thoughts about violence and, because of that, it does not put in jeopardy the deep experience of meaning, significance, and sense that individuals experience as a result of their membership in the group (e.g., Haslam et al., 2020; Swann et al., 2012). Given that ingroup members tend to be trusted more and to exert more influence than people who do not belong to the group (e.g., Brewer, 1999; Cialdini & Griskevicious, 2019), even instilling moral beliefs that blatantly oppose the use of violence among some individuals who belong to violent extremist groups may be a force of positive social change for other group members.

Visceral Responsibility and Identity Fusion

Chapter 5 includes two cross-sectional studies and one experiment exploring whether strongly fused individuals are so extraordinarily willing to engage on extreme self-sacrifices to promote the interest of the individuals (Study 1a) or groups (Studies 1b and 2) with which they

feel one because they experience visceral responsibility towards them. They also explore the role of shared responsibility in the attenuation of the consequences of fusion.

Studies 1a and 1b showed that, irrespectively of the target of fusion (the romantic partner, Study 1a, or the country, Study 1b), identity fusion is positively related to feelings of visceral responsibility and willingness to fight and die for the fusion target, and that the relationship between identity fusion and willingness to fight and die is apparently mediated by feelings of visceral responsibility. Study 2 goes a step further offering causal evidence in support of the former mechanism and demonstrating the moderating role played by shared feelings of visceral responsibility. The results of this study show that learning that other group members display low levels of visceral responsibility diminishes personal feelings of visceral responsibility and attenuates the impact of identity fusion on willingness to fight and die. The results also showed that identity fusion is positively related to willingness to fight and die even when individuals learn that feelings of responsibility are not shared; and that the relationship between both variables is mediated by personal feelings of visceral responsibility in both experimental conditions.

As the studies presented in the preceding chapters, the studies included in Chapter 5 make important contributions to the theoretical and practical domain. Theoretically, our studies advance identity fusion and tie it to the ethics of care in three different ways. First, by demonstrating that strongly fused individuals are heavily inclined to fight and die for the target of fusion because they feel viscerally responsible to protect and promote its wellbeing and interests, our studies reveal a novel mediational mechanism not researched before and highlight that, without need for further abstract reasoning or justification, identity fusion may give rise to automatic, instinctive, and impulsive feelings of visceral responsibility that push individuals to make sacrificial behaviours, as posited by some descriptions of responsibility or natural care (e.g., Noddings, 2003; Tronto, 1993). Second, our studies also contribute to previous research

on the factors that moderate the relationship between identity fusion and proneness to engage in extreme self-sacrifices, showing that strongly fused individuals are less inclined to fight and die for the group when they discover that feelings of visceral responsibility are not widely shared. These results lend indirect support to the idea that highly fused individuals see themselves as embedded in caring relationships and that some degree of shared trust and solidarity are needed to maintain a high level of self-sacrifice, as has been pointed by several care theorists as well (e.g., Tronto, 2013; van Nistelrooij, 2015). Third, by showing that strongly fused individuals are willing to fight and die even when feelings of responsibility are not shared, our studies reveal that the relationship between identity fusion and extreme self-sacrifice does not depend on strict reciprocity or hard mutualism, also as posited by some care ethicists (e.g., Slote, 2007; van Nistelrooij, 2015).

At a practical level, our studies suggest several strategies to prevent the risk of violence related to moral parochialism. For instance, when individuals are already fused with violent groups, the use of persuasive messages challenging the assumption that feelings of visceral responsibility are shared may attenuate their willingness to engage in violence. This strategy alone might not suffice to impede aggression, but it has the advantage of being relatively cheap. In general, it may also be useful to develop intervention programs aimed at the development of fusion with humanity, because that can counter our tendency to fight against each other in circumstances of conflict and inspire us to work together to reach mutually benefiting outcomes (Talaifar & Swann, 2018). One of the most promising paths for achieving this goal seems to be the practice of loving-kindness and compassion meditation (Adair et al., 2018; Hutcheson et al., 2008; Kang et al., 2014, 2015).

Limitations

The research included in this thesis has several limitations. First, our ***studies on the relationship between admiration and identity fusion*** have three main deficiencies. The first

limitation is related to the fact that we tested our model with imprisoned Jihadists and the general population and, as such, we cannot guarantee that the results also apply to other groups, such as different types of terrorists or social activists. The second limitation relates to the finding that the effects of admiration on identity fusion and extreme behaviours do not vary as a function of the number of individuals towards whom admiration is felt. Although identity fusion is based on the development of relational ties and this fact could explain our results, abundant research suggests that the efficacy of social influence attempts augments as the number of people who hold a certain point of view or engage in a particular act increases (e.g., Cialdini & Griskevicius, 2019; Cialdini & Sagarin, 2005). Given that, a single study showing that the number of agents towards whom admiration is experienced does not exert any effect may not be enough to ensure that we encountered a robust result (Goh et al., 2016). Lastly, the third limitation is associated to the fact that feeling admiration is a mood enhancer (e.g., Algoe & Haidt, 2009; Onu et al., 2016) and that positive mood motivates individuals to act prosocially (e.g., Baron, 1997; North et al, 2004; Salovey et al., 1991). Despite numerous studies have manifested that the effects of admiration cannot be explained by other positive emotional states, we did not measure positive emotions and, consequently, we are not able to discard the possibility that the effects that we found in our studies could be partially or totally attributable to them.

Second, our ***studies on the moderating role of moral beliefs about the justifiability of violence*** present two remarkable limitations. The first limitation stems from the fact that we did not include manipulation checks to find out if participants really imagined that the fusion target was under a serious threat, although this is a minor concern because strongly fused individuals are extremely sensitive to information depicting the group as being under threat (e.g., Sheikh et al., 2016; Vázquez et al., 2018) and we mentioned the existence of threat explicitly in all the items included in our measures. The second limitation derives from the fact

that our studies used a correlational approach and they do not allow to establish causality. By definition, moral convictions are highly resistant to social influence attempts (e.g., Aramovich et al., 2011; Malo, 2021; Skitka & Morgan, 2014); and we have been trying to manipulate moral beliefs for other unrelated studies several years without success. Because of that, we renounced to conduct experiments.

Third, the main limitation of the ***studies on the mediating role of feelings of visceral responsibility*** is that they do not provide direct evidence to support the idea that identity fusion causes feelings of visceral responsibility and not the other way around, although theoretical reasons incline us to believe this.

Lastly, a common limitation of the studies focusing on moral beliefs and visceral responsibility is that we conducted all the studies online, with samples composed by students at an open university - i.e., UNED - and their acquaintances. Despite such individuals are characterized by having a higher sociodemographic variability than the students at face-to-face universities, our methodology might rise some doubts about the external validity and the generalizability of the findings to more ecologically valid settings and special samples, such as violent extremist offenders, social activists, volunteers, or combatants. Also, although all these studies measure constructs that relate to actual pro-group behaviour (i.e., behavioural intentions), we did not include measures of real pro-group acts in any of them.

Future Lines of Research

Our studies suggest some future lines of research as well. In the ***narrative review article***, we summarize some of the avenues that could be pursued by scientists to further develop identity fusion theory and dismiss some of the misconceptions that we identified during our reading of the publications within the field. Regarding the *development of the theory*, we pointed out that conducting research with samples of adolescents from 12 to 15 years old would be particularly useful to better understand how individuals move from a state

of protofusion to a mature state of fusion and gain a better comprehension of the ontogenetic development of this visceral feeling of oneness. We also signalled that examining the relative influence of shared positive and negative experiences and shared core values on identity fusion and its correlates could help us disentangle the relevance of each of these factors. Lastly, we posited that more research should be conducted to explore the underlying mechanisms involved when people are fused with entities different from the group, because there may be considerable variation regarding the antecedents, consequences, mediators, and moderators when the target of fusion differs. Concerning the *misconceptions about identity fusion*, we indicated that examining the phenomena that can be predicted by group identification and not by identity fusion would be helpful to further distinguish both constructs; and that comparing identity fusion to other predictors of extreme behaviour would help us to establish the relative weight of each of them.

Since the moment at which we published the theoretical review, several scientists have walked some steps in various of the directions that we suggested. For instance, Paredes et al. (2021) showed that the moderating role of the extremity of the situation applies to fusion with values as well as to fusion with groups; we demonstrate that feelings of visceral responsibility mediate the relationship between fusion with individuals and groups and extreme pro-group behaviours (Chapter 5); Martel el al. (2021) contrasted the predictive power of identity fusion to that of moral mandates and sacred values; Wolfowick et al. (2021) pitted fusion against a plethora of risk factors for violent radicalization; and White et al. (2021) found that group identification with hard-core football fans predicts outgroup anxiety and prejudice, whereas identity fusion predicts outgroup hostility.

In the ***article focusing on the study of the causal impact of admiration***, we highlighted some avenues of research that may be useful to solve the main limitations of our studies. Specifically, conducting more field research to try to replicate our model with additional

special samples of interest; performing some studies to examine whether relational ties explain the lack of impact of the number of individuals towards whom admiration is felt on identity fusion and extreme self-sacrifices; and conducting studies including measures of mood and other positive emotions and controlling for these variables to show that our results are genuinely due to admiration. Complementarily to that, other studies could examine whether feelings of gratitude - an other-oriented emotion that promotes improved relationships with benefactors (e.g., Algoe & Haidt, 2009) - towards the outgroup can attenuate the tendency of strongly fused individuals to engage in violence to defend and protect their group and motivate them to pursue peaceful means to solve antagonisms; or whether feeling hate towards the outgroup increases willingness to self-sacrifice for the ingroup and/or harm the outgroup by augmenting the extent to which individuals feel fused to their group (e.g., Halperin, 2016).

In the ***article centred on the moderating role of moral beliefs about the justifiability of violence***, we suggested several studies useful to complement our research. Among others, we mentioned that some studies should examine whether the effect of personal moral beliefs that we discovered in our studies also extends to other ideological factors, such as the norms of the group or its narratives (e.g., Elnakouri et al., 2018; Ginges & Shackleford, 2018); that other studies should delve into the boundary conditions under which moral beliefs against the use of violence exert their effects and determine if some of the mechanisms related to moral disengagement (Bandura, 1990), like the displacement of responsibility or the dehumanization of the other, attenuate, or even eliminate, their impact; and that some research should examine the processes lying at the core of the development of proscriptive beliefs opposed to violence among the strongly fused. In relation to the last point, we signalled that the experience of some emotions, like disgust or anger (Wineski & Skitka, 2017), may be important for this process of moralization, but other factors could play a relevant role too. For instance, Tomasello (2016) has asserted that moral norms are internalized when individuals assume that they have

contributed to their development and experience some feelings of co-authorship; and several authors have posited that moral beliefs turn essential as behavioural guides when individuals have the conviction that they are indispensable to reach socio-moral progress (e.g., Gray, 2003; 2014; Malo, 2021)

In the ***article focusing on the examination of the mediating role played by feelings of visceral responsibility***, we also mentioned some lines of research that could complement our studies, like exploring whether the mediational mechanism that we discovered generalises to identity fusion with mindless entities and examining the specific emotions that might give rise to feelings of visceral responsibility, such as sympathetic concern or empathy (e.g., Slote, 2007). Following the ideas of several ethicists of care (e.g., Ginters, 1976; van Nislelrooj, 2015), it could also be interesting to explore if strongly fused people are prone to engage in *expressionate acts* towards the target of fusion, that is, in behaviours that are not aimed to directly promote the well-being of the fusion target but that are valuable *per se* because they express its deep significance for the person who cares, such as bringing flowers to the grave of a deceased person or honouring the image of a group that has long disappeared. This last research has the potential to tie identity fusion to the ethics of care in another novel and unexplored way.

Conclusions

Humans have been talking about individuals who make marvellous and terrible personal sacrifices for the groups that are dear to them since immemorial times. Their deeds awake our passions making us laugh from hope and weep from despair. The consequences of their acts are so disparate that we feel automatically inclined to believe that there cannot be anything in common between them, and yet identity fusion research has demonstrated that their behaviours are frequently motivated by the same psychological state: a feeling of connection to the group so deep that changes their self and makes them prone to incur in great risks and

personal sacrifices to protect and defend the group and each of its members. Here, we merge identity fusion theory with research on morality and show that admiring such individuals might instigate identity fusion with group or values and augment the will to go to the extremes for them. We also show that strongly fused people are disposed to use violent means to promote the interests of the group when they think that violence is morally justifiable, but not when they do not believe so. Lastly, we show that the moral machinery that explains their inclination to engage in extreme prosocial actions is built upon feelings of visceral responsibility based on care. We wish that our research helps to contemplate identity fusion under a more correct light as well as to raise awareness about the utmost importance of promoting fusion with peaceful groups and providing individuals with a strong moral education that impedes them from dismissing the negative consequences of violence.

Conclusiones

Los humanos han estado hablando de individuos que realizan sacrificios personales maravillosos y terribles por los grupos que aprecian desde tiempos inmemoriales. Sus hazañas despiertan nuestras pasiones y nos hacer reír de esperanza y llorar de desesperación. Las consecuencias de sus actos son tan dispares que nos sentimos automáticamente inclinados a creer que no puede haber nada en común entre ellos y, sin embargo, la investigación sobre la fusión de la identidad ha demostrado que sus conductas están frecuentemente motivadas por el mismo estado psicológico: un sentimiento de conexión con el grupo tan profundo que cambia su yo y les hace propensos a incurrir en grandes riesgos y sacrificios personales para proteger y defender al grupo y a cada uno de sus miembros. Aquí, combinamos la teoría de la fusión de la identidad con la investigación sobre moralidad y mostramos que admirar a dichos individuos podría instigar la fusión de la identidad con grupo o valores y aumentar la disposición a llegar a los extremos por ellos. También mostramos que las personas fuertemente fusionadas están dispuestas a usar medios violentos para promover los intereses del grupo cuando piensan que

la violencia es moralmente legítima, pero no cuando no lo creen. Por último, mostramos que la maquinaria moral que explica su inclinación a implicarse en actos prosociales extremos está construida sobre sentimientos de responsabilidad visceral basados en el cuidado. Deseamos que nuestra investigación ayude a contemplar la fusión de la identidad bajo una luz más correcta, así como a aumentar la consciencia de la importancia crucial que tiene promover la fusión con grupos pacíficos y dotar a los individuos de una fuerte educación moral que les impida desestimar las consecuencias negativas de la violencia.

REFERENCES

Adair, K. C., Fredrikson, B. L., Castro-Schilo, L., Kim, S., & Sidberry, S. (2018). Present with you: Does cultivated mindfulness predict greater social connection through gains in decentering and reductions in negative emotions. *Mindfulness, 9*(3), 737-749. https://doi.org/10.1007/s12671-017-0811-1

Admiration. Oxford English Dictionary. https://www.oed.com/view/Entry/2566. Accessed May 28, 2020

Agerström, J., Carlsson, R., Nicklasson, L., & Guntell, L. (2016). Using descriptive social norms to increase charitable giving: The power of local norms. *Journal of Economic Psychology, 52,* 147-153. https://doi.org/10.1016/j.joep.2015.12.007

Algoe, S.B., & Haidt, J. (2009) Witnessing excellence in action: the 'other-praising' emotions of elevation, gratitude, and admiration. *Journal of Positive Psychology, 4(2),* 105-127. https://doi.org/10.1080/17439760802650519

Apps, M. A. J., McKay, R., Azevedo, R. T., Whitehouse, H., & Tsakiris, M. (2018). Not on my team: medial prefrontal cortex responses to ingroup fusion and unfair monetary divisions. *Brain and Behavior, 8*(8), 1-13. https://doi.org/10.1002/brb3.1030

Aquino, K., McFerran, B., & Laven, M. (2011) Moral identity and the experience of moral elevation in response to acts of uncommon goodness. *Journal of Personality and Social Psychology, 100*(4), 703-718. https://doi.org/10.1037/a0022540

Aramovich, N. P., Lytle, B. L., & Skitka, L. J. (2011). Opposing torture: Moral conviction and resistance to majority influence. *Social Influence, 7*(1), 21-34. https://doi.org/10.1080/15534510.2011.640199

Ashokkumar, A., Galaif, M., & Swann, W. B. Jr. (2019) Tribalism can corrupt: Why people denounce or protect immoral group members. *Journal of Experimental Social Psychology, 85*, 103874. https://doi.org/10.1016/j.jesp.2019.103874

Ashour, O. (2010). Online de-radicalization? Countering violent extremist narratives: Message, messenger and media strategy. *Perspectives on Terrorism, 4*(6), 15-19. https://www.jstor.org/stable/26298491

Atran, S. (2010). *Talking to the enemy: Violent extremism, sacred values, and what it means to be human.* Penguin.

Atran, S. (2019) The Islamic State's Lingering Legacy among Young Men from the Mosul Area. Combating Terrorism Center at West Point. http://ctc.usma.edu/islamic-states-lingering-legacy-among-young-men-mosul-area/. Published January 17, 2019. Accessed May 14, 2020

Atran, S. (2020). Psychology of transnational terrorism and extreme political conflict. *Annual Review of Psychology.* https://doi.org/10.1146/annurev-psych-010419-050800

Atran, S., Sheikh, H., & Gómez, A. (2014). Devoted actors sacrifice for close comrades and a sacred cause. *Proceedings of the National Academy of Sciences, 111*(50), 17702-17703. https://doi.org/10.1073/pnas.1420474111

Atran, S., Waziri, H., Gómez, Á., Sheikh, H., López-Rodríguez, L., Rogan, C., & Davis, R. (2018). The Islamic State's lingering legacy among young men from the Mosul area. *CTC Sentinel*, *11*(4), 15-22.

Babinska, M., & Bilewicz, M. (2018), Self-sacrifice for in-group's history: A diachronic perspective. *Behavioral and Brain Sciences, 41*, e194. https://doi.org/10.1017/S0140525X1801796

Bairak, F. (2019). *Morality as a moderator of the relationship between group identification and ingroup favoritism* [Unpublished Master's thesis]. Baskent University.

Bandura, A. (1990). Selective activation and disengagement of moral control. *Journal of Social Issues, 46*(1), 27-46. https://doi.org/10.1111/j.1540-4560.1990.tb00270.x

Baron, R. A. (1997). The sweet smell of…helping: Effects of pleasant and ambient fragances on prosocial behavior in shopping malls. *Personality and Social Psychology Bulletin, 23*, 498-503. https://doi.org/10.1177/0146167297235005

Batson, C. D. (2019). *A scientific search for altruism. Do we care only about ourselves?* Oxford University Press.

Baumeister, R. F., Bratlavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323-370. https://doi.org/10.1037/1089-2680.5.4.323

Baumeister, R. F., & Leary, M. R. (1995) The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497-529. https://doi.org/10.1037/0033-2909.117.3.497

Bélanger, J. J. (2017). The Rise and Fall of Violent Extremism: The Science Behind Community-Based Interventions. In C. Köpetz. & A. Fishbach (Eds.), *The motivation-cognition interface; from the lab to the real world* (pp. 170-195). Routledge.

Bélanger, J. J., Caouette, J., Sharvit, K., & Dugas, M. (2014) The psychology of martyrdom: Making the ultimate sacrifice in the name of a cause. *Journal of Personality and Social Psychology, 107*(3), 494-515. https://doi.org/10.1037/a0036855

Bélanger, J. J., Moyano, M., Muhammad, H., Richardson, L., Lafrenière, M. A. K., McCaffery, P., ...Nociti, N. (2019). Radicalization leading to violence: A test of the 3N model. *Frontiers in Psychiatry*, *10*, 42. https://doi.org/10.3389/ftsyt.2019.00042

Bélanger, J. J., Schumpe, B. M. M., Menon, B., Ng, J. C., & Nociti, N. (2018) Self-Sacrifice for a Cause: A Review and an Integrative Model. In V. Zeigler-Hill V. & T. K. Shackelford (Eds.), *The SAGE Handbook of Personality and Individual Differences: Volume II: Origins of Personality and Individual Differences*. https://doi.org/10.4135/9781526451200.n25

Berkowitz, L. (1972). Social norms, feelings, and other factors affecting helping and altruism. *Advances in Experimental Social Psychology, 6*, 63-108. https://doi.org/10.1016/S0065-2601(08)60025-8

Besta, T., Gómez, A., & Vázquez, A. (2014). Readiness to deny group's wrongdoing and willingness to fight for its members: The role of Poles' identity fusion with the country and religious group. *Current Issues in Personality Psychology, 2*(1), 49-55. https://doi.org/10.5114/CIPP.2014.43101

Besta, T., Jaśkiewicz, M., Kosakowska-Berezecka, N., Lawendowski, R., & Zawadzka, A. M. (2018) What do I gain from joining crowds? Does self-expansion help to explain the relationship between identity fusion, group efficacy and collective action? *European Journal of Social Psychology, 48*, O152– O167. https://doi.org/10.1002/ejsp.2332

Bjørgo, T. (2011) Dreams and disillusionment: engagement in and disengagement from militant extremist groups. *Crime, Law and Social Change, 55*(4), 277-285. https://doi.org/10.1007/s10611-011-9282-9

Bjørgo, T., & Horgan, J. (2009) Leaving Terrorism Behind: Individual and Collective Disengagement. Routledge.

Bortolini, T. S., Newson, M., Natividade, J. C., Vázquez, A., & Gómez, A. (2018). Identity fusion predicts endorsement of pro-group behaviors targeting nationality, religion or football in Brazilian samples. *British Journal of Social Psychology, 57*(2), 346-366. https://doi.org/10.1111/bjso.12235

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate. *Journal of Social Issues, 55*(3), 429-444. https://doi.org/10.1111/0022-4537.00126

Buhrmester, M. D., Burnham, D., Johnson, D. D. P., Curry, O. S., Mcdonald, D. V., & Whitehouse, H. (2018). How moments become movements: Shared outrage, group

cohesion and the lion that went viral. *Frontiers in Ecology and Evolution, 6*, 54. https://doi.org/10.3389/fevo.2018.00054.

Buhrmester, M. D., Fraser, W. T., Lanman, J. A., Whitehouse, H., & Swann, W. B. Jr. (2014). When terror hits home: Identity fused Americans who saw Boston bombing victims as "family" provided aid. *Self and Identity, 14*(3), 253-270. https://doi.org/10.1080/15298868.2014.992465

Buhrmester, M. D., Gómez, A., Brooks, M. L., Morales, J. F., Fernández, S., & Swann, W. B. (2012). My Groups Fate Is My Fate: Identity-Fused Americans and Spaniards Link Personal Life Quality to Outcome of 2008 Elections. *Basic Applied Social Psychology 34*(6), 527-533. https://doi.org/10.1080/01973533.2012.732825

Buhrmester, M. D., Newson, M., Vázquez, A., Hattori, W. T., & Whitehouse, H. (2018). Winning at any cost: Identity fusion, group essence, and maximizing ingroup advantage. *Self and Identity, 17*(5), 500-516. https://doi.org/10.1080/15298868.2018.1452788

Buhrmester, M. D., & Swann, W. B. Jr. (2015). Identity fusion. In R. Scott and S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences.* John Wiley and Sons. https://doi.org/10.10002/9781118900772.etrds0172.

Carnes, N. C., & Lickel, B. (2018). Moral binding: How emotions, convictions, and identity fusion shape progroup behavior. *Self and Identity, 17*(5), 549-573. https://doi.org/10.1080/15298868.2018.1451362

Chinchilla, J., Vázquez, A., & Gómez, A. (2021). Identity fusion predicts violent behavior when it is morally justifiable. *The Journal of Social Psychology.* https://doi.org/10.1080/00224545.2021.1948813

Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591-621. https://doi.org/10.1146/annurev.psych.55.090902.142015

Cialdini, R. B., & Griskevicius, V. (2019) Social influence. In Finkel E. J. & Baumeister R. F. (Eds.), *Advanced social psychology. The state of the science* (2nd Ed.). Oxford University Press.

Cialdini, R. B., & Sagarin, B. (2005) Principles of interpersonal influence. In Brock T. C. & Green M. C. (Eds.) *Persuasion: Psychological insights and perspectives.* Sage Publications.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp., 151-192). McGraw-Hill.

Clark, M. S. (1984). Record keeping in two types of relationships. *Journal of Personality and Social Psychology, 47*(3), 549-557. https://doi.org/10.1037/0022-3514.47.3.549

Clark, M. S., Dubash, P., & Mills, J. (1998). Interest in another's consideration of one's needs. *Journal of Experimental Social Psychology, 34*(3), 246-264. https://doi.org/10.1006/jesp.1998.1352

Cox, K. S. (2010). Elevation predicts domain-specific volunteerism 3 months later. *Journal of Positive Psychology, 5*(5), 333-341. https://doi.org/10.1080/17439760.2010.507468

Crimston, D., & Hornsey, M. J. (2018). Extreme self-sacrifice beyond fusion: moral expansiveness and the special case of allyship. *Behavioral and Brain Sciences, 41*, e198. https://doi.org/10.1017/S0140525X18001620.

Curry, O. S. (2016). Morality as cooperation: A problem-centered approach. In R. D. Hansen (Ed.), *The evolution of morality* (pp.27-52). Springer International Publishing. https://doi.org/10.1007/978-3-319-19671-8_2

Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 6O societies. *Current Antropology, 60*(1), 47-49. https://doi.org/10.1086/701578

Dessalles, J. L. (2018). Self-sacrifice as a social signal. *Behavioral and Brain Sciences, 41*, e200. https://doi.org/10.1017/S0140525X18001590.

Deutsch, M., Gerald, H. B. (1955). A study on normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*(3), 629-636. https:/doi.org/10.1037/h0046408

Earp, B. D., Mc Loughlin, K. L., Monrad, J. T., Clark, M. S., & Crockett, M. J. (2021). *How social relationships shape moral judgment.* Manuscript submitted for publication. https://doi.org/10.31234/osf.io/e7cgq

Ellemers, N. (2017). *Morality and the regulation of social behavior: Groups as moral anchors.* Routledge. https://doi.org/10.4324/9781315661322

Elnakouri, A., McGregor, I, & Grossman, I. (2018). The importance of environmental threats and ideology in explaining extreme self-sacrifice. *Behavioral and Brain Sciences, 41,* e201. https://doi.org/10.1017/S01405X18001723

Englander, Z. A., Haidt, J., & Morris, J. P. (2012). Neural basis of moral elevation demonstrated through inter-subject synchronization of cortical activity during free-viewing. *PLoS ONE, 7*(6), e39384. https://doi.org/10.1371/journal.pone.0039384

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: a three-factor theory of anthropomorphism. *Psychological Review, 114*(4), 864-886. https://doi.org/10.1037/0033-295X.114.4.864

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

Ferguson, N., & McAuley, J. W. (2020). Staying engaged in terrorism: Narrative accounts of sustaining participation in violent extremism. *Frontiers in Psychology.* https://doi.org/10.3389/fpsyg.2020.01338

Fiske, F. (1992). The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review, 99*(4), 689-723. https://10.1037/0033-295x.99.4.689

Fiske, A., P., & Rai, T. S. (2015). *Virtuous violence*. Cambridge University Press. https://doi.org/10.1017/CBO9781316104668

Fredman, L. A., Bastian, B., & Swann, W. B. Jr. (2017). God or country? Fusion with Judaism predicts desire for retaliation following Palestinian stabbing intifada. *Social Psychologycal and Personality Science, 8*(8), 882-887. https://doi.org/101177/1948550617693059

Fredman, L.A., Buhrmester, M. D., Gómez, A., Fraser, W. T., Talaifar, S., Brannon, S. M., & Swann, W. B. Jr. (2015). Identity fusion, extreme pro-group behavior, and the path to defusion. *Social and Personality Psychology Compass, 9*(9), 468–480. https://doi.org/10.1111/spc3.12193

Gaviria, E., Ferreira, C., Martínez, M., & Whitehouse, H. (2015). Identity and the developmental origins of fusion: An exploratory approach. *International Journal of Social Psychology, 30*(3), 531-562. https://doi.org/10.1080/02134748.2015.10650088

Gilligan, C. (1982). *In a different voice*. Harvard University Press.

Ginges, J., & Shackleford, C. (2018). Self-sacrifice for a cause: The role of ideas and beliefs in motivating human conflict. *Brain and Behavioral Sciences, 41,* e203. https://doi.org/10.1017/S0140525X18001693

Ginters, R. (1976). *Die Ausdruckhandlung. Eine Untersuchung ihrer sittlichen Bedeutsamkeit.* Patmos.

Goh, J. K., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass, 10*(10), 535-549. https:doi.org/10.1111/spc312267

Gómez, A. (2005, January). Identity Fusion. In *Society of Personality and Social Psychology Conference.* Conference celebrated at Utah, United States of America.

Gómez, A., Bélanger, J. J., Chinchilla, J., Vázquez, A., Schumpe, B. M., Nisa, C. F., & Chiclana, S. (2021). Admiration for Islamist groups encourages self-sacrifice though identity fusion. *Humanities and Social Sciences Communications, 8*, 54. https://doi.org. 10.1057/s41599-021-00734-9

Gómez, A., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & Swann W. B. Jr. (2011). On the nature of identity fusion: Insights into the construct and a new measure. *Journal of Personality and Social Psychology, 100*(5), 918-933. hptts://doi.org/10.1037/a0022642

Gómez, A., Chinchilla, J., Vázquez, A., López-Rodríguez, L., Paredes, B., & Martínez, M. (2020). Recent advances, misconceptions, untested assumptions, and future research agenda for identity fusion theory. *Social and Personality Psychology Compass,* e12531. https://doi.org/10.1111/spc3.12531

Gómez, A., López-Rodríguez, L., Sheikh, H., Ginges, J., Wilson, L., Waziri, H., Vázquez, A., Davis, R., & Atran, S. (2017). The devoted actor's will to fight and the spiritual dimension of human conflict. *Nature Human Behavior 1*(9), 673-679. https://doi:10.1038/s41562-017-0193-3

Gómez, A., Martínez, M., Martel, F. A., López-Rodríguez, L., Vázquez, A., Chinchilla, J., Paredes, B., Hettiarachchi, M., Nafees, H., & Swann, W. B. (2021). Why people enter and embrace violent groups. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.614657

Gómez, A., Morales, J. F., Hart, S., Vázquez, A., & Swann, W. B. Jr. (2011). Rejected and excluded forevermore, but even more devoted: Irrevocable ostracism intensifies loyalty to the group among identity-fused persons. *Personality and Social Psychology Bulletin, 37*(12), 1574-1586. https://doi.org/10.1177/0146167211414580

Gómez, A., & Vázquez, A. (2015). The power of 'feeling one' with a group: identity fusion and extreme pro-group behaviors. *International Journal of Social Psychology, 30*, 481-511. https://doi.org/10.1080/02134748.2015.1065089

Gómez, A., Vázquez, A., López-Rodríguez, L., Talaifar, S., Martínez, M., Buhrmester, M. D., & Swann, W. B. Jr. (2019). Why people abandon groups: Degrading relational vs. collective ties uniquely impacts identity fusion and identification. *Journal of Experimental Social Psychology, 85*, 103853. https://doi.org/10.1016/j.jesp.2019.103853

Gray, J. (2003). *Straw dogs: Thoughts on humans and other animals*. Granta Books.

Gray, J. (2014). *The silence of animals: On progress and other modern myths.* Farrar Strauss & Giroux.

Greene, J. (2013). *Moral tribes: Emotion, reason, and the gap between us and them.* Penguin Press.

Haidt, J. (2003) The Positive emotion of elevation. *Prevention and Treatment, 3*(1). https://doi.org10.1037/1522-3736.3.1.33c

Halperin, E. (2016). *Emotions in conflict: Inhibitors and facilitators of peace making*. Routledge.

Hamid, N., Pretus. C, Atran, S., Crockett, M. J., Ginges, J., Sheikh, H.,… Vilarroya, O. (2019). Neuroimaging 'will to fight' for sacred values: an empirical case study with supporters of an Al Qaeda associate. *Royal Society Open Science, 6*(6), 181585. https://doi.org/10.1098/rsos.181585

Haslam, S. A., Haslam, C., Cruwys, T., Jetten, J., Bantley, S. V., Fong, & Steffens, N. K. (2021). Social identity makes group-based connection possible: Implications for loneliness and mental health. *Current Opinion in Psychology, 24*(43), 161-165. https://doi.org/10.1016/jcopsyc.2021.07.13

Hawkins, M. A. (2019). The effect of activity identity fusion on negative consumer behavior. *Psychology and Marketing, 36*(4), 395-409. https://doi.org/10.1002/mar.21186

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd Ed.). Guildford Press.

Heger, A. K., & Gaertner, L. (2018). Testing the identity synergy principle: Identity fusion promotes self and group sacrifice. *Self and Identity, 17*(5), 487-499. https://10.108'/15298868.2017.1422538

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*(3),165-196. https://doi.org/10.1016/s1090-5138(00)00071-4

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*, 1820-1300. https://doi.org/10.1037//0003-066x.52.12.1280

Higgins, E. T., Roney, C. J. R., Crowe, E., & Hymes, C. (1994). Ideal versus ought prediction for approach and avoidance: Distinct regulatory systems. *Journal of Personality and Social Psychology, 66*, 276-286. https://doi.org/10.1037//0022-3514.66.2.276

Hogg, M. A., Cooper-Shaw, L., & Holzworth, D. W. (1993). Group prototypicality and depersonalized attraction in small interactive groups. *Personality and Social Psychology Bulletin, 19*(4), 452-465. https://doi.org/10.1177/0146167293194010

Hogg, M. A., & Hardie, E. A. (1991). Social attraction, personal attraction, and self-categorization. A field study. *Personality and Social Psychology Bulletin, 17*(2), 175-180. https://doi.org/10.1177/014616729101700209

Hogg, M. A., Hardie, E. A., & Reynolds, K. J. (1995). Prototypical similarity, self-categorization, and depersonalized attraction: A perspective on group cohesiveness. *European Journal of Social Psychology, 25*(2), 159-177. https://doi.org/10.1002/ejsp.2420250204

Hume, D. (1739/2000) *A treatise of human nature*. Oxford University Press.

Hutcheson, C. A., Seppala, E. M., & Gross, J. J. (2008). Loving-kindness meditation increases social connectedness. *Emotion, 8*(5), 720-724. https://doi.org/10.1037/a0013237

Immordino-Yang, M. H., McColl, A., Damasio, H., & Damasio, A. (2009) Neural correlates of admiration and compassion. *Proceedings for the National Academy of Sciences, 106*(19), 8021-8026. https://doi.org/10.1073/pnas.0810363106

Jacobson, M. (2009). Terrorist drop-outs: One way of promoting a counter-narrative. *Perspectives on Terrorism, 3*(2), 12-17.

Janoff-Bulman, R., Sheikh, S., & Hepp, S. (2009). Proscriptive versus prescriptive morality: two faces of moral regulation. *Journal of Personality and Social Psychology, 96*(3), 521-537. https://doi.org/10.1037/a0013779

Jiménez, J., Gómez, A., Buhrmester, M. D., Vázquez, A., Whitehouse, H., & Swann, W. B. Jr. (2016). The Dynamic Identity Fusion Index: A new continuous measure of identity fusion for web-based questionnaires. *Social Science Computer Review, 34*(2), 215-228. https://doi.org/10.1177/0894439314566178

Jong, J., Whitehouse, H., Kavanagh, C., & Lane, J. (2015). Shared negative experiences lead to identity fusion via reflection. *PLoS ONE, 10*(12): e0145611. https://doi.org/10.1371/journal.pone.0145611

Joo, M., & Park, S. W. (2017). Effect of identity fusion on decision to make extreme sacrifices in romantic relationships: The moderating role of impulsiveness. *British Journal of Social Psychology, 56*, 819-827. https://doi.org/10.1111/bjso.12218

Kang, Y., Gray, J. F., & Dovidio, J. F. (2014). The nondiscriminating heart: lovingkindness meditation training decreases implicit intergroup bias. *Journal of Experimental Psychology General, 143*(3), 1306-1313. https://doi.org/10.1037/a0034150

Kang, Y., Gray, J. F., & Dovidio, J. F. (2015). The head and the heart: effects of understanding and experience lovingkindness on attitudes toward the self and others. *Mindfulness, 6*(5), 1063-1070. https://doi.org/10.1007/s12671-014-0355-6

Kant (1788/2011). *Kritik der praktischen Vernunft*. Klassiker Auslegen.

Kapitány, R., Kavanagh, C., Buhrmester, M. D., Newson, M., & Whitehouse, H. (2019). Ritual, identity fusion, and the inauguration of President Trump: A pseudo-experiment of ritual modes theory. *Self and Identity, 19*(3), 293-323. https://doi.org/10.1080/15298868.2019.1578686

Kavanagh, C. M., Jong, J., McKay, R., & Whitehouse, H. (2018). Positive experiences of high arousal martial arts rituals are linked to identity fusion and costly pro-group actions. *European Journal of Social Psychology, 49*(3), 461-481. https://doi.org/10.1002/ejsp.2514

Keltner D., & Haidt, J. (2003) Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion, 17*(2), 297-314. hptts://doi.org/10.1080/02699930302297

Kiper, J., & Sosis, R. (2018). Toward a more comprehensive theory of self-sacrificial violence. *Behavioral and Brain Sciences, 41,* e206. https://doi.org/10.1017/S0140525X18001802

Kishna, A., & Kim, S. (2021). Exploring the dynamics between brand investment, customer investment, brand identification, and brand identity fusion. *Journal of Business Research, 137*, 267-277. https://doi.org/10.1016/j.jbusres.2021.08.039

Kruglanski, A. W., Bélanger, J., & Gunaratna, R. (2019). The three pillars of radicalization: needs, narratives and networks. Oxford University Press. https://doi.org/10.1093/oso/9780190851125.001.0001.

Kruglanski, A. W., Gelfand, M. J., Bélanger, J. J., Gunaratna, R., & Hettiararchchi, M. (2014). Deradicalizing the Liberation Tigers of Tamil Eelam (LTTE): Some Preliminary Findings. In Silke, A. (Ed.), *Prisons, Terrorism and Extremism: Critical Issues in Management, Radicalization and Reform*. Routledge.

Kruglanski, A. W., Webber, D., & Koehler, D. (2019) *The radical's journey: How German Neo-Nazis voyaged to the edge and back.* Oxford Scholarship online. https://doi.org/10.1093/oso/9780190851095.001.0001

Kunst, J. R., Boos, B., Kimel, S. Y., Obaidi, M., Shani, M., & Thomsen, L. (2018). Engaging in extreme activism in support of others' political struggles: The role of politically motivated fusion with outgroups. *PLoS ONE, 13*(1), e0190639. https://doi.org/10.1371/journal.pone.019639

Kunst, J. R., Dovidio, J. F., & Thomsen, L. (2019). Fusion with political leaders predicts willingness to persecute immigrants and political opponents. *Nature Human Behavior, 3*(11), 1180-1189. https://doi.org/10.1038/s41562-019-0708-1

Lankford, A. (2018). Strength in numbers: A survival strategy that helps explain social bonding and commitment. *Behavioral and Brain Sciences, 41*, e208. https://doi.org/10.1017/S0140525X18001553

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

Lemay, E. P., Overall, N. C., & Clark, M. S. (2010). Experiences and interpersonal consequences of hurt and anger. *Journal of Personality and Social Psychology, 103*(6), 982-1006. https://doi.org.10.1037/a0030064

Lench, H. C., Taylor, A. B., & Bench, S. W. (2014). An alternative approach to analysis of mental states in experimental social cognition research. *Behavior Research Methods, 46,* 215-228. https://doi.org/10.3758/s13428-013-0351-0

Leonard, D. J., Moons, W., Mackie, D., & Smith, E. R. (2011): "We are mad as hell and we're not going to take it anymore": Anger self-stereotyping and collective action. *Group Processes & Intergroup Relations, 14*(1), 99-111. https://doi.org/10.1177/1368430210373779

Lobato, R. M., & Sainz, M. (2019). On the way to fusion through the pilgrims' route: Factors that maintain identity fusion in collective rituals. *Group Processes & Intergroup Relations, 23*(4), 502-518. https://doi.org/10.1177/1368430219849690

Lobato, R. M., & Sainz, M. (2021). Collective rituals in times of the COVID-19 quarantine: The relationship between collective applause and identity fusion. *Studia Psychologica, 63*(3), 279-290. https://doi.org/10.311577/sp.2021.03.827

Louis, W. R., McGarty, Thomas, E. F., Amiot, C. E., & Moghaddam, F. M. The power of norms to sway fused group members. *Behavioral and Brain Sciences, 41*, e209. https//doi.org/10.1017/S0140525X18001644

Maio, G. R. (2019). *The psychology of human values*. Psychology Press.

Malo, J. (2021). *Los peligros de la moralidad. Por qué la moral es una amenaza para las sociedades del siglo XXI.* Deusto.

Martel, F. A., Buhrmester, M., Gómez, A., Vázquez, A., & Swann, W. B. Jr. (2021). Why true believers make the ultimate sacrifice: Sacred values, moral convictions, or identity

fusion? *Frontiers in Psychology, 12*, 779120. https://doi.org/10.3389/fpsyg/2021.779120

Milla, M. N., Putra, I. E., & Umam, A. N. (2019). Stories from Jihadists: Significance, identity and radicalization trough the call for Jihad. *Peace and Conflict: Journal of Peace Psychology, 25*(2), 111-121. https://doi.org/10.1037/pac0000371

Misch, A., Ferguson, G., & Durnham, Y. (2018). Temporal dynamic of partisan identity fusion and prosociality during the 2006 U.S. Presidential Election. *Self and Identity, 17*(5), 531-548. https://doi.org/10.1080/15298868.2018.1430063

Newson, M. (2017). Football, fan violence, and identity fusion. *International Review of the Sociology of Sport, 54*(4), 431-444. https://doi.org/10.1177/1012690217731293

Newson, M. (2021). High and highly bonded: Fused football fans who use cocaine are most likely to be aggressive toward rivals. *International Journal of Drug Policy, 93*, 103263. https://doi.org/10.1016/j.drugpo.2021.103263

Newson. M., Bortolini, T., Buhrmester, M., da Silva, S. R., Queiroga da Aquino, J. N., & Whitehouse, H. (2018). Brazil's football Warriors: Social bonding and inter-group violence. *Evolution and Human Behavior, 39*(6), 675-683. https://doi.org/10.1016/j.evolhumbehav.2018.06.010

Newson, M., Buhrmester, M., & Whitehouse, H. (2016). Explaining lifelong loyalty: The role of self-shaping identity fusion and self-shaping group events. *PLoS ONE, 11*(8), e0160427. https://doi.org/10.1371/journal.pone.0160427

Noddings, N. (2003). *Caring: A relational approach to ethics and moral education* (updated). UCLA.

Nook, E. C., Ong, D. C., Morelli, S. A., Mitchell, J. P., & Zaki, J. (2016). Prosocial conformity: Prosocial norms generalize across behavior and empathy. *Personality and Social Psychology Bulletin, 42*(8), 1045-1062. https://doi.org/10.1177/0146167216649932

North, A. C., Tarrant, M., & Hargreaves, D. J. (2004). Effects of music on helping behavior: A field study. *Environmental Behavior, 36*, 266-275. https://doi.org/10.1177/0013916503256263

Nyamutata, C. (2020). Young Terrorists or Child Soldiers? ISIS Children, International Law and Victimhood. *Journal of Conflict and Security Law*. https://doi.org/10.1093/jcsl/krz034

Olivola, C. Y. (2018). The motivation to sacrifice for a cause reflects a basic cognitive bias. *Behavioral and Brain Sciences, 41*, e212. https://doi.org/10.1017/S0140525X18001747

Onu, D., Kessler, T., & Smith, J. R. (2016). Admiration: A Conceptual Review. *Emotion Review 8*(3), 218-230. https://doi.org/10.1177/1754073915610438

Páez, D., Rimé, B., Basabe, N., Wlodarcsyk, A., & Zumeta, L. (2015). Psychological effects of perceived synchrony in collective gatherings. *Journal of Personality and Social Psychology, 108*(5), 711-729. https://doi.org/10.1037/pspi0000014

Paredes, B., Briñol, P., & Gómez, A. (2018). Identity fusion leads to willingness to fight and die for the group: The moderating impact of being informed of the reasons behind other members' self-sacrifice. *Self and Identity, 17*(5), 517-530. https://doi.org/10.1080/15298868.2017.1419503

Paredes, B., Briñol, P., Petty, R. E., & Gómez, A. (2021). Increasing predictive validity of identity fusion leading to sacrifice by considering the extremity of the situation. *European Journal of Social Psychology, 51*(2), 239-250. https://doi.org/10.1002/EJSP/2727

Paredes, P., Santos, D., Briñol, P., Gómez, A., & Petty, R. E. (2019). The role of meta-cognitive certainty on the relationship between identity fusion and endorsement of extreme pro-

group behavior. *Self and Identity, 19*(1), 1-20. https://doi.org/10.1080/15298868.2019.1681498

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. (2005) Prosocial Behavior: Multilevel Perspectives. *Annual Review of Psychology, 56*(1), 365-392. https://doi.org/10.1146/annurev.psych.56.091103.070141

Pohling, R., & Diessner, R. (2016). Moral Elevation and Moral Beauty: A Review of the Empirical Literature. *Review of General Psychology, 20*(4), 412-425. https://doi.org/10.1037/gpr0000089

Pretus, C., Hamid, N., Sheikh, H., Ginges, J., Tobeña, A., Davis, R.,… Atran, S. (2018). Neural and Behavioral Correlates of Sacred Values and Vulnerability to Violent Extremism. *Frontiers in Psychology, 9*, 2462. https://doi.org/10.3389/fpsyg.2018.02462

Pryor, C., Perfors, A., & Howe, P. D. L. (2019). Even arbitrary norms influence moral decision-making. *Nature Human Behavior, 3*, 57-62. https://doi.org/10.1038/s41562-018-0489-y

Purzycki, B. G., & Lang, M. (2019). Identity fusion, outgroup relations, and sacrifice: A cross cultural test. *Cognition, 186*, 1-6. https://doi.org/10.1016/j.cognition.2019.01.015

Reese, E., & Whitehouse, H. (2021). The development of identity fusion. *Perspectives on Psychological Science, 16*(6), 1398-1411. https://doi.org/10.1177/1745691620968761

Reicher, A., Haslam, S. A., & Rath, R. (2008). Making a virtue of evil: A five-step social identity model of the development of collective hate. *Social and Personality Psychology Compass, 2*(3), 1313-1344. https://doi.org/10.1111/j.1751-9004.2008.00113.x

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*(4), 296-320. https://doi.org/10.1207/S15327957PSPR0504_2

Saleem, M., Anderson, C. A., & Barlett, C. P. (2015). Assessing Helping and Hurting Behaviors Through the Tangram Help/Hurt Task. *Personality and Social Psychology Bulletin, 41*(10), 1345-1362. https://doi.org/10.1177/0146167215594348

Salovey, R., Mayer, J. D., & Rosenhan, D. L. (1991). Mood and helping: Mood as motivator of helping and helping as regulator of mood. In Clark M. (Ed.), *Review of personality and social psychology. Prosocial behavior* (Vol. 12). Sage.

Schindler, I., Zink, V., Windrich, J., & Menninghaus, W. (2013). Admiration and adoration: Their different ways of showing and shaping who we are. *Cognition and Emotion, 27*(1), 85-118. https://doi.org/10.1080/02699931.2012.698253

Schnall, S., Roper, J., & Fessler, D. M. (2010). Elevation leads to altruistic behavior. *Psychological Science, 21*(3), 315-320. https://doi.org/10.1177/0956797609359882

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science, 8*(4), 379-386. https://doi.org/10.1177/1948550617715068

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18*(5), 429-434. https://doi.org/10.1111/j.1467-9280.2007.01917.x

Segal, K., Jong, J., & Halberstadt, J. (2018), The fusing powers of natural disasters: An experimental study. *Self and Identity, 17*(5), 574-586. doi: 10.1080/15298868.2018.1458645.

Sheikh, H., Gómez, A., & Atran, S. (2016). Empirical evidence for the devoted actor model. *Current Anthropology, 57*(13), S204-S209. https://doi.org/10.1086/686221

Siromahov, M., Buhrmester, M., & McKay, R. (2020). Beliefs in national continuity are related to essentialist thinking and to perceptions of the nation as family. *Nations and Nationalism, 26*, 845-863

Skitka, L. J., & Morgan, G. S. (2014). The social and political implications of moral conviction. In H. Lavine (Ed.), *Advances in Political Psychology*. Bloomsbury Academic Press.

Slote, M. (2007). *The ethics of care and empathy*. Routledge.

Smith, R. H. (2000). Assimilative and Contrastive Emotional Reactions to Upward and Downward Social Comparisons. In Suls, J. & Wheeler, L. (Eds.), *Handbook of Social Comparison. Theory and Research*. Kluwer Academic Publishers. https://doi:10.1007/978-1-46154237-7_10

Speckhard, A., Shajkovci, A., & Bodo, A. (2018): Fighting ISIS on Facebook: Breaking the ISIS brand counter-narratives project. *International Center for the Study of Violent Extremism*, 50-66.

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89*(6), 845-851. https://doi.org/10.1037/0022-3514.89.6.845

Stellar, J. E., Gordon, A. M., Piff, P. K. et al (2017). Self-transcendent emotions and their social functions: Compassion, gratitude, and awe bind us to others through prosociality. *Emotion Review, 9*(3), 200-207. https://doi.org/10.1177/1754073916684557

Swann, W. B., Jr., & Buhrmester, M. (2015). Identity fusion. *Current Directions in Psychological Science*, 24, 52–57. https://doi.org/10.1177/0963721414551363

Swann, W. B. Jr., Buhrmester, M. D., Gómez, A., Jetten, J., Bastian, B., Vázquez, A., …, & Zhang, A. (2014). What makes a group worth dying for? Identity fusion fosters perception of familial ties, promoting self-sacrifice. *Journal of Personality and Social Psychology, 106*(6), 912-926. https://doi.org/10.1037/a0036089

Swann, W. B. Jr., Gómez, A., Buhrmester, M. D., López-Rodríguez, L., Jiménez, J., & Vázquez, A. (2014). Contemplating the ultimate sacrifice: Identity fusion channels pro-

group affect, cognition, and moral decision making. *Journal of Personality and Social Psychology, 106*(5), 713-727. https://doi.org/10-1037/a0035809

Swann, W. B. Jr., Gómez, A., Dovidio, J. F., Hart, S., & Jetten, J. (2010). Dying and killing for one's group. Fusion moderates responses to intergroup versions of the trolley problem. *Psychological Science, 21*(8), 1176-1183. https://doi.org/10.1177/0956797610376656

Swann, W. B. Jr., Gómez, S., Huici, C., Morales, F., & Hixon, J. G. (2010). Identity fusion and self-sacrifice: Arousal as catalyst of pro-group fighting, dying, and helping behavior. *Journal of Personality and Social Psychology, 99*(5), 824-841. https://doi.org/10.1037/a0020014

Swann, W. B., Jr., Gómez, A., Seyle, D. C., Morales, J. F., & Huici, C. (2009). Identity fusion: The interplay of personal and social identities in extreme group behavior. *Journal of Personality and Social Psychology, 95*(5), 995-1011. https://doi.org/10.1037/a0013668

Swann, W. B. Jr., Gómez, A., Vázquez, A., Gillamón, A., Segovia, S., & Carrillo, B. (2015). Fusion with the cross-gender group predicts sex reassignment surgery. *Archives of Sexual Behavior, 44*(5), 1313-1318. https://doi.org/10.1007/s10508-014-0470-4

Swann, W. B. Jr., Jetten, J., Gómez, A., Whitehouse, H., & Bastian, B. (2012). When group membership gets personal: A theory of identity fusion. *Psychological Review, 119*(3), 441-456. https://doi.org/10.1037/a0028589

Swann, W. B. Jr., Seyle, D. C., Gómez, A., & Morales, F. J. (2005, January). Identity fusion. In *The social psychology winter conference*. Conference celebrated at Utah, United States of America.

Sweetman, J., Spears, R., Livingstone, A. G., & Manstead A. S. (2013). Admiration regulates social hierarchy: Antecedents, dispositions, and effects on intergroup behavior. *Journal*

of Experimental Social Psychology, 49(3), 534-542. https://doi.org/10.1016/j.jesp.2012.10.007

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Brooks-Cole.

Talaifar, S., Ashokkumar, A. Pennebaker, J. W., Medrano, F. N., Yeager, D. S., & Swann, W. B. Jr. (2020). A new pathway to university retention? Identity fusion with university predicts retention independently of grades. *Social Personality and Psychology Science, 12*(1), 108-117. https://doi.org/10.1117.1948550619894995

Talaifar, S., & Swann, W. B. Jr. (2018). Identity fusion: The union of personal and social selves. In P. J. Ivanhoe, O. J. Flanagan, V. S. Harrison, H. Sarkissian, & E. Schwitzgebel (Eds.), *The oneness hypothesis: Beyond the boundary of the self.* Columbia University Press.

Tapley, M., & Clubb, G. (2019). *The role of formers in countering violent extremism.* International Centre for Counter-Terrorism.

Tasuji, T., Reese, E., van Mulukom, V., & Whitehouse, H. (2020). Band of mothers: Childbirth as a female bondig experience. *PLoS ONE, 15*(10), e0240175. https://doi.org/10.1371/journalpone.0240175

Thomsen, L., & Fiske, A. P. (2018). Communal sharing/identity fusion does not require reflection on episodic memory of shared experience of trauma -and usually generate kindness. *Behavioral and Brain Sciences, 41*, e220. https://doi.org/10.1017/S0140525X18001784

Time (2019) Greta Thunberg: TIME's Person of the Year 2019. https://time.com/person-of-the-year-2019-greta-

thunberg/?utm_source=twitter&utm_medium=social&utm_campaign=person-of-the-year&utm_term=_&linkId=78763203. Accessed May 14, 2020

Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.

Tornero, E., Sánchez-Romera, J. F., Morosoli, J. J., Vázquez, A., Gómez, A., & Ordoñana, J. R. (2017). Altruistic behavior among twins. *Human Nature, 29*(1), 1-12. https://doi.org/10.1007/s12110-017-9304-0

Tronto, J. C. (2013, March). *Democratic caring and global responsibilities for care* [Paper presentation]. Annual Meeting of the Western Political Science Association, Hollywood, USA. https://www.wpsanet.org/papers/docs/Tronto%20WPSA%20paper%202013.pdf

Tronto, J. C. (1993). *Moral boundaries. A political argument for an ethic of care.* Routledge.

Tronto, J. C. (1998). An ethic of care. *Generations: Journal of the American Society of Aging, 22*(3), 15-20.

UN. Security Council Committee (2019) Letter dated 15 July 2019 from the Chair of the Security Council Committee Established pursuant to Resolutions 1267 (1999), 1989 (2011) and 2253 (2015) concerning Islamic State in Iraq and the Levant (Da'esh), Al-Qaida and Associated Individuals, Groups, Undertakings and Entities addressed to the President of the Security Council. United Nations Digital Library. http//digitallibrary.un.org/record/3813209?ln=en#record-files-collapse-header. Accessed May 28, 2020

van De Ven, N., Archer, A., & Engelen, B. (2018). More important and surprising actions of a moral exemplar trigger stronger admiration and inspiration. *The Journal of Social Psychology, 159*(4), 383-397. https://doi.org/10.1080/00224545.2018.1498317

van Nistelrooij, I. (2015). *Sacrifice: A care-ethical reappraisal of sacrifice and self-sacrifice*. Peeters.

Vázquez, A., Gómez, A., Ordoñana, J. R., & Paredes, B. (2015). From interpersonal to extended fusion: relationships between fusion with siblings and fusion with the country. *International Journal of Social Psychology, 30*(3), 512-530. https://doi.org/10.1080/02134748.2015.1093755

Vázquez, A., Gómez, A., Ordoñana, J. R., Swann, W. B. Jr., & Whitehouse, H. (2017). Sharing genes fosters identity fusion and altruism. *Self and Identity, 16*(6), 684-702. https://doi.org/10.1080/15298868.2017.1296887

Vázquez, A., Gómez, A., & Swann, W. B. Jr. (2017). Do historic threats to the group diminish identity fusion and its correlates. *Self and Identity, 16*(4), 480-503. https://doi.org/10.1080/15298868.2016.1272485

Vázquez, A., López-Rodríguez, L., Martínez, M., Atran, S., & Gómez, Á. (2020). Threat enhances aggressive inclinations among devoted actors via increase in their relative physical formidability. *Personality and Social Psychology Bulletin*, 1-15. https://doi.org/10.1177/014616722090746

Vázquez, A., Ordoñana, J. R., Whitehouse, H., & Gómez, A. (2019). Why die for my sibling? The positive association between identity fusion and imagined loss with endorsement of self-sacrifice. *International Journal of Social Psychology, 34*(3), 413-438. https://doi.org/10.1080/02134748.2019.1639343

Vianello, M., Galliani, E. M., & Haidt, J. (2010). Elevation at work: The effects of leaders' moral excellence. *Journal of Positive Psychology, 5*(5), 390-411. https:// doi.org/10.1080/17439760.2010.516764

Vignoles, V. (2019). Identity: Personal AND social. In K. Deaux and M. Snyder (Eds.), *Oxford handbook of personality and social psychology* (Second Edition). Oxford University Press.

Vyver, J. V. D., & Abrams, D. (2015). Testing the prosocial effectiveness of the prototypical moral emotions: Elevation increases benevolent behaviors and outrage increases justice behaviors. *Journal of Experimental Social Psychology, 58*, 23-33. https://doi.org/10.1016/j.jesp.2014.12.005

Walsh, C. M., & Neff, L. A. (2018). We're better when we blend: The benefits of couple identity fusion. *Self and Identity, 17*(5), 587-603. https://doi.org/10.1080/15298868.2018.1430062

Webber, D., Chernikova, M., Kruglanski, A. W., Gelfand, M. J., Hettiarachchi, M., Gunaratna, R., Lafreniere, M. A., & Belanger, J. J. (2018). Deradicalizing detained terrorists. *Poitical Psychology*, *39*(3), 539-556. https://doi.org/10.1111/pops.12428

Webber D, & Kruglanski A. W. (2016). Psychological factors in radicalization: A "3 N" approach. In G. LaFree & J. D. Freilich (Eds), *The Handbook of the Criminology of Terrorism* (pp. 33–46.). Wiley.

Weisbuch, M., & Ambady, N. (2008). Affective divergence: Automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology, 95*, 1063-1079. https://doi.org/10.1037/a0011993

Whitaker, R. M., Colombo, G. B., & Dunham, Y. (2021). The evolution of strongly-held group identities trough agent based cooperation. *Scientific Reports, 11*, 12071. https://doi.org/10.1038/s41598-021-91333-w

White, F. A., Newson, M., Verrelli, S., & Whitehouse, H. (2021). Pathways to prejudice and outgroup hostility: Group alignment and intergroup conflict among football fans. *Journal of Applied Social Psychology.* https://doi.org/10.1111/jasp.12773

Whitehouse, H., (2018). Dying for the group: Towards a general theory of extreme self-sacrifice. *Behavioral and Brain Sciences, 41*, e192, 1-62. https://doi.org/10.1017/S0140525X18000249

Whitehouse, H., & Fitzgerald, R. (2020). Fusion and reform: The potential for identity fusion to reduce recidivism and improve reintegration. *Anthropology in Action, 27*(1), 1-13. https://doi.org/10.3167/aia.2020.270101

Whitehouse, H. Jong, J., Buhrmester, M. D., Gómez, A., Bastian, B., Kavanagh, C. M., Newson, M., Matthews, M., Lanman, J. A., McKay, R., & Gavrilets, S. (2017). The evolution of cooperation via shared dysphoric experiences. *Scientific Reports, 7*, 44292. https://doi.org/10.1038/srep44292

Whitehouse, H., McQuinn, B., Buhrmester, M., & Swann, W. B. Jr. (2014). Brothers in arms: Libyan revolutionaries bound like family. *Proceedings for the National Academy of Sciences, 111*(50), 17783-17785. https://doi.org/10.1073/pnas.1416284111

Wiessner (2018). Considering selection pressures for identity fusion and self-sacrifice in small scale societies. *Behavioral and Brain Sciences,* e220. https://doi.org/10.1017/S0140525X18001681

Wineski, D. C., & Skitka, L. J. (2017). Moralization through moral schock: Exploring emotional antecedents to moral conviction. *Personality and Social Psychology Bulletin, 43*(2), 139-150. https://doi.org/10.1177/0146167216676479

Wolfowicz, M., Litmanovitz, Y., Weisburd, D., & Hasisi, B. (2021) Cognitive and behavioral radicalization: A systematic review of the putative risk and protective factors. *Campbell Systematic Reviews, 17*(3)*,* e1174. https://doi.org/10.1002/cl2.1174

Zaal, M. P., C. Ståhl, T., Ellemers, N., & Derks, B. (2011). By any means necessary: The effects of regulatory focus and moral conviction on hostile and benevolent forms of collective action. *British Journal of Social Psychology, 50*(4)*,* 670-689. hppts://doi.org/10.1111/j.2044-8309.2011.02069.x

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2018). Cognitive underpinnings of nationalistic ideology in the context of Brexit. *Proceedings of the National Academy of*

Sciences of the United States of America, 115(19), E4532-E4540. https://doi.org/10.1073/pnas.1708960115

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2019). Cognitive inflexibility predicts extreme attitudes. *Frontiers in Psychology,* 10, 989. https://doi.org/10.3389/fpsyg.2019.00989

Zumeta, L., Basabe, N., Wlodarczyk, A., Bobowik, M., & Páez, D. (2016). Shared flow and positive collective gatherings. *Anales de Psicología*, *32*(3), 717-727. https://doi.org/10.6018/analesps.32.3.261651

ANNEX A

‚Empirical Studies Published in the Present PhD Thesis

DOI: 10.1111/spc3.12531

ARTICLE

WILEY

Recent advances, misconceptions, untested assumptions, and future research agenda for identity fusion theory

Ángel Gómez[1,2] | Juana Chinchilla[2] | Alexandra Vázquez[1,2] | Lucía López-Rodríguez[1,3] | Borja Paredes[4] | Mercedes Martínez[2]

[1]ARTIS International, St. Michaels, Maryland

[2]Faculty of Psychology, Social and Organizational Psychology Department, Universidad Nacional de Educación a Distancia, Madrid, Spain

[3]Psychology Department, Universidad de Almería, Almería, Spain

[4]Faculty of Sciences of Information, Theory and Analisys of Comunication Department, Universidad Complutense de Madrid, Madrid, Spain

Correspondence
Ángel Gómez, Faculty of Psychology, Social and Organizational Psychology Department, Universidad Nacional de Educación a Distancia, Madrid, Spain.
Email: agomez@psi.uned.es

Funding information
Ministerio de Ciencia, Innovación y Universidades from the Spanish Government, Grant/Award Number: RTI2018-093550-B-100

Abstract

Just a decade ago, two psychologists, Swann, and Gómez, developed a new theoretical framework to explain extreme pro-group behaviors: *identity fusion theory*. *Identity fusion* refers to a visceral feeling of oneness with a group that motivates individuals to do extraordinary self-sacrifices on behalf of the group or each of its members. Since the formulation of the theory, interdisciplinary researchers of the five continents have conducted dozens of studies on identity fusion, both in laboratory and field settings. Research has deepened into the causes, consequences, underlying mechanisms, and applications of identity fusion. The development of fusion-based research has been steadfast and very prolific. Hence, the first section of the current manuscript includes an updated overview of this fast growing literature. This increasing interest for the theory has, however, been accompanied by a series of misconceptions and untested research assumptions, which we address in the second and third sections of the paper, concluding with a final section suggesting a future research agenda. Our aim is to help those interested in knowing more about identity fusion or about the causal mechanisms that lead individuals to risk their life and personal well-being for a group discarding common misconceptions as well as formulating more precise and nuanced hypotheses for future research.

Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-021-00734-9 OPEN

Admiration for Islamist groups encourages self-sacrifice through identity fusion

Á. Gómez[1,2,5], J. J. Bélanger[3,5✉], J. Chinchilla[2], A. Vázquez[1,2], B. M. Schumpe[3], C. F. Nisa[3] & S. Chiclana[4]

The psychological mechanisms that lead terrorists to make costly sacrifices for their ideological convictions are of great theoretical and practical importance. We investigate two key components of this process: (1) the feeling of admiration toward ingroup members making costly self-sacrifices for their ideological group, and (2) identity fusion with religion. Data collected in 27 Spanish prisons reveal that jihadists' admiration toward members of radical Islamist groups amplifies their willingness to engage in costly sacrifices for religion in prison. This effect is produced because admiration toward radical Islamist groups has a binding effect, increasing identity fusion with religion. Five additional experiments provide causal and behavioural evidence for this model. By showing that admiration for ingroup members increases identity fusion, which in turn makes individuals prone to engage in costly pro-group behaviours, we provide insights into the emotional machineries of radicalization and open new avenues for prevention strategies to strengthen public safety.

[1] Artis International, Washington, DC, USA. [2] Universidad Nacional de Educación a Distancia, Madrid, Spain. [3] New York University Abu Dhabi, Abu Dhabi, United Arab Emirates. [4] Secretaría General de Instituciones Penitenciarias Ministerio del Interior, Madrid, Spain. [5] These authors contributed equally: Á. Gómez, J.J. Bélanger. ✉email: jocelyn.belanger@nyu.edu

THE JOURNAL OF SOCIAL PSYCHOLOGY
https://doi.org/10.1080/00224545.2021.1948813

Routledge
Taylor & Francis Group

Identity fusion predicts violent pro-group behavior when it is morally justifiable

Juana Chinchilla, Alexandra Vazquez, and Ángel Gómez

UNED

ABSTRACT

Identity fusion is a visceral feeling of oneness with a group that predicts extreme pro-group behaviors. However, the effects of fusion depending on the nature of such behaviors -violent versus nonviolent- and the factors that may incline strongly fused individuals to display them still remain unexplored. To fill this gap, we performed two correlational studies in which we examined whether moral beliefs regarding the justifiability of violence moderate the relationship between fusion with the family (Study 1), or with the country (Study 2), and willingness to engage in violent and nonviolent pro-group acts. Results showed that strongly fused participants were more willing to act violently than weakly fused participants, but only when their beliefs in the moral justifiability of violence were high. In contrast, their willingness to engage in nonviolent acts was not influenced by moral beliefs.

ARTICLE HISTORY
Received 8 January 2021
Accepted 12 June 2021

KEYWORDS
Identity fusion; self-sacrificial behavior; violence; morality

Identity fusion is a visceral feeling of oneness with a group that predicts extreme pro-group behaviors. Recent research has demonstrated that this feeling of extraordinary connection can also be established to another individual, an animal, an object, or an activity (Gómez et al., 2020). Although dozens of studies conducted in five continents and in different contexts have consistently shown that identity fusion predicts costly sacrifices and willingness to fight and die for the group among others, and preliminary research has used identity fusion as an approach to explain football violence (Newson, 2017), no empirical research up to date has explored whether the nature of such pro-group behavior (violent vs. nonviolent) affects the predictive character of fusion, and whether beliefs in the moral justifiability of violence moderate the relationship between fusion and pro-group actions. Identifying the variables that determine whether strongly fused individuals pursue the goals of the group through violent versus nonviolent activities would provide important insights into the conditions under which identity fusion has negative or positive social consequences (Swann & Buhrmester, 2015), and that is the main aim of this research.

Research on identity fusion originated in the aftermath of the 9/11 attacks in New York, and the Madrid train bombings, as an attempt to explain why some terrorists and members of other violent groups engage in extreme self-sacrificial behaviors (Swann et al., 2009). Identity fusion theory postulates that this type of acts might result from a psychological state in which individuals feel so deeply entrenched into the group and emotionally committed to it that they are willing to promote its interests and that of its members even when that comes at a high cost to the self (Gómez et al., 2020; Swann et al., 2014). Consequently, a considerable amount of research has included measures of willingness to fight and die and to engage in costly pro-group sacrifices and has found that fusion is positively related to them (e.g., Gómez, Brooks et al., 2011; Gómez et al., 2017; Gómez, Morales et al., 2011; Swann, Gómez, Huici et al., 2010; Swann et al., 2009). Other research has also shown that the

CONTACT Juana Chinchilla juanachinchilla@me.com Facultad de Psicología, Departamento de Psicología Social y de las Organizaciones, 28040, Madrid, (Spain)
Supplemental data for this article can be accessed on the publisher's website

Received: 27 September 2021 | Accepted: 2 February 2022

DOI: 10.1111/bjso.12526

ARTICLE

Strongly fused individuals feel viscerally responsible to self-sacrifice

Juana Chinchilla[1] | **Alexandra Vázquez**[1,2] | **Ángel Gómez**[1,2]

[1]Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain

[2]ARTIS International, St. Michaels, Maryland, USA

Correspondence
Juana Chinchilla, Dpto. Psicología Social y de las Organizaciones, C/ Juan del Rosal, Nº 10 (Ciudad Universitaria), CP. 28040, Madrid, Spain.
Email: jchinchilla@psi.uned.es

Funding information
Spanish Ministry of Science, Innovation and Universities, Grant/Award Numbers: RTI2018-098576-A-I00, RTI2018-093550-B-I00

Abstract

Identity fusion is a visceral feeling of oneness that predicts extreme behaviour on behalf of the target of fusion. We propose that strongly fused individuals are characterized by feelings of visceral responsibility towards such target – unconditional, instinctive, and impulsive drive to care, protect and promote its well-being and interests – that motivates them to self-sacrifice. Two studies offered initial support when the target of fusion is an individual or a group (Studies 1a-1b). A final study added causal evidence that strongly fused learning that most ingroup members did not feel visceral responsibility towards the group expressed less willingness to self-sacrifice than those learning that ingroup members display high levels of visceral responsibility (Study 2). These findings offer novel evidence for the mechanisms underlying the effects of fusion on extreme behaviour on behalf of the target of fusion and the attenuation of its consequences.

KEYWORDS

care ethics, extreme behaviour, identity fusion, prosocial behaviour

INTRODUCTION

> If I take on the other´s reality as a possibility and begin to feel its reality, I feel also that *I must* act accordingly, that is, I am impelled to act as in my own behalf, but in behalf of the other.
>
> Nel Noddings (2003; cursives are ours)

ANNEX B

Supplementary Materials for Chapter 3: Admiration for Islamist Groups Encourages Self-sacrifice through Identity Fusion

Study 1 (N = 71): Admiration Toward Radical Islamist Groups Increases Identity Fusion with Religion and Willingness to Engage in Costly Sacrifices among Imprisoned Jihadists

Method and Design

Measures

Costly sacrifices for religion. We measured costly sacrifices for religion using a five-item scale developed for the study ("If it was necessary, I would be willing to give up communications with my family or important people outside prison to defend my religion", "If it was necessary, I would we willing to stop taking part in activities that make me feel good to defend my religion", "If it was necessary, I would be willing to lost commodities into prison (e.g. and individual room, my own clothes, my own hygienic stuff) to defend my religion," "If it was necessary, I would be willing to give up my prison income to defend my religion" and "If it was necessary, I would be willing to move to a prison further away from my family to defend my religion").

Study 2 (*N* = 152): Feelings of Admiration Increase Perceived Cost of Self-sacrifice, Identity Fusion, and Willingness to Fight and Die

Results

Preliminary Analyses

Feelings of admiration. An ANOVA on feelings of admiration toward the ingroup member showed a significant effect of condition, $F(1, 150) = 296.85$, $p < .001$, $\eta^2_p = .66$. Participants felt more admiration toward the ingroup member in the admiration condition ($M = 4.80$, $SD = 0.87$) than in the typical condition ($M = 2.15$, $SD = 1.02$).

Main Analyses

Fusion with other groups. Due to technical issues with the download of the dynamic measures, we could only collect and analyse data from 94 participants. A MANOVA on fusion with the other groups revealed that our manipulation did not influence participants fusion with family, $F(1, 92) = 0.35$, $p = .56$ $\eta^2_p = .00$, France, $F(1, 92) = 0.46$, $p = .497$, $\eta^2_p = .00$, Portugal, $F(1, 92) = 0.40$, $p = .529$ $\eta^2_p = .00$, or Italy, $F(1, 92) = 0.00$, $p = .962$, $\eta^2_p = .00$.

Study 3 ($N = 231$): Feelings of Admiration toward an Individual versus a Group

Results

Preliminary Analyses

Feelings of admiration. An ANOVA on feelings of admiration showed that the main effect of level of admiration was significant, $F(1, 227) = 1897.35$, $p < .001$, $\eta^2_p = .89$. Participants felt more admiration in the admiration conditions ($M = 4.85$, $SD = 0.98$) than in the no admiration conditions ($M = 0.25$, $SD = 0.55$). The effects of target of admiration, $F(1, 227) = 1.57$, $p = .212$, $\eta^2_p = .01$, and the interactive effect, $F(1, 227) = 0.70$, $p = .403$, $\eta^2_p = .00$, were not significant.

Main Analyses

We performed a 2 (no admiration vs. admiration) x 2 (individual vs. group) multivariate analysis of variance (MANOVA) on cost of sacrifice, fusion with the group, and willingness to fight and die.

Cost of sacrifice. The analyses showed that the effect of the interaction was not significant, $F(1, 227) = 0.28$, $p = .598$, $\eta^2_p = .00$.

Identity Fusion. Neither the main effect of target of admiration, $F(1, 227) = 0.005$, $p =$

.946, η^2_p = .00, nor the effect of the interaction were significant, $F(1, 227) = 0.002$, $p = .969$, η^2_p = .00.

Willingness to fight and die. The effects of target of admiration, $F(1, 227) = 1.06$, $p = .305$, η^2_p = .005, and the interactive effect, $F(1, 227) = 0.00$, $p = .981$, η^2_p = .00, were not significant.

Fusion with other groups. Because participants had technical issues with downloading the dynamic measures, we could only collect and analyse the data from 135 participants. A 2 X 2 factorial MANOVA on fusion with other groups indicated that the main effect of level of admiration was not significant on fusion with family, F(1, 131) = 0.21, p = .644, η^2_p = .00, France, F(1, 131) = 1.57, p = .212, η^2_p = .01, Portugal, F(1, 131) = 1.52, p = .219, η^2_p = .01, or Italy, F(1, 131) = 0.21, p = .647, η^2_p = .00. The main effect of target of admiration was not significant on fusion with family, F(1, 131) = 0.18, p = .671, η^2_p = .001, France, F(1, 131) = 2.93, p = .089, η^2_p = .02, Portugal, F(1, 131) = 0.46, p = .497, η^2_p = .00, or Italy, F(1, 131) = 0.09, p = .770, η^2_p = .00. Finally, the interaction of level of admiration by target of admiration was not significant on fusion with family, F(1,131) = 0.64, p = .424, η^2_p = .01, France, F(1, 131) = 0.489, p = .48, η^2_p = .00, Portugal, F(1,131) = 0.01, p = .937, η^2_p = .00, or Italy, F(1, 131) = 0.11, p = .739, η^2_p = .00.

Study 4 (N = 253): Feelings of Admiration towards an Ingroup versus an Outgroup Member

Results

Preliminary Analyses

Feelings of admiration. A 2 (no admiration vs. admiration) x 2 (ingroup vs. outgroup) ANOVA on feelings of admiration showed that the main effect of level of admiration was

significant, $F(1, 249) = 735.44$, $p < .001$, $\eta^2_p = .75$. Participants felt more admiration in the admiration conditions ($M = 4.12$, $SD = 1.31$) than in the no admiration conditions ($M = 0.34$, $SD = 0.82$). The effect of group, $F(1, 249) = 1.64$, $p = .201$, $\eta^2_p = .01$, and the interaction, $F(1, 249) = 1.17$, $p = .281$, $\eta^2_p = .00$, were not significant.

Main Analysis

We performed a 2 (no admiration vs. admiration) x 2 (ingroup vs. outgroup) multivariate analysis of variance (MANOVA) on cost of sacrifice, fusion with the group, and willingness to fight and die.

Cost of sacrifice. The analysis showed a significant main effect of level of admiration, $F(1, 249) = 84.69$, $p < .001$, $\eta^2_p = .25$ (admiration conditions: $M = 3.71$, $SD = 1.32$; no admiration conditions: $M = 2.35$, $SD = 1.00$). The main effect of group was not significant, $F(1, 249) = 0.15$, $p = .701$, $\eta^2_p = .00$ (ingroup conditions: $M = 3.11$, $SD = 1.40$; outgroup conditions: $M = 3.03$, $SD = 1.32$).

Identity Fusion. The analysis showed a significant main effect of level of admiration, $F(1, 249) = 15.21$, $p < .001$, $\eta^2_p = .06$ (admiration conditions: $M = 2.26$, $SD = 1.27$; no admiration conditions: $M = 1.68$, $SD = 1.03$). The main effect of group was not significant, $F(1, 249) = 1.78$, $p = .183$, $\eta^2_p = .01$ (ingroup conditions: $M = 3.11$, $SD = 1.40$; outgroup conditions: $M = 3.03$, $SD = 1.32$).

Willingness to fight and die. Results showed a significant main effect of level of admiration, $F(1, 249) = 11.04$, $p = .001$, $\eta^2_p = .04$ (admiration conditions: $M = 1.93$, $SD = 1.04$; no admiration conditions: $M = 1.58$, $SD = 0.50$). The main effect of group was not significant, $F(1, 249) = 1.17$, $p = .281$, $\eta^2_p = .01$ (ingroup conditions: $M = 1.83$, $SD = 1.02$; outgroup conditions: $M = 1.70$, $SD = 0.62$).

Fusion with other groups. Due to technical issues with the download of the dynamic

measures, we could only collect and analyse the data from 145 participants. A 2 X 2 factorial MANOVA on fusion with other groups indicated that the main effect of level of admiration was not significant on fusion with family, $F(1, 141) = 1.84$, $p = .177$, $\eta^2_p = .01$, France, $F(1, 141) = 0.14$, $p = .704$, $\eta^2_p = .00$, Portugal, $F(1,141) = 1.10$, $p = .296$, $\eta^2_p = .01$, or Italy, $F(1, 141) = 0.07$, $p = .793$, $\eta^2_p = .00$. The main effect of group was not significant on fusion with family, $F(1, 141) = 0.59$, $p = .443$, $\eta^2_p = .00$, France, $F(1, 141) = 1.53$, $p = .218$, $\eta^2_p = .01$, Portugal, $F(1, 141) = 1.94$, $p = .166$, $\eta^2_p = .01$, or Italy, $F(1, 141) = 0.94$, $p = .333$, $\eta^2_p = .01$. Lastly, the effect of the interaction was not significant on fusion with family, $F(1, 141) = 0.85$, $p = .358$, $\eta^2_p = .01$, France, $F(1,141) = 0.42$, $p = .517$, $\eta^2_p = .00$, Portugal, $F(1, 141) = 0.06$, $p = .804$, $\eta^2_p = .00$, or Italy, $F(1,141) = 0.002$, $p = .963$, $\eta^2_p = .00$.

Study 5 ($N = 241$): Feelings of Admiration Due to Self-sacrifice versus Personal Qualities

Results

Preliminary Analyses

Feelings of admiration. A 2 (no admiration vs. admiration) x 2 (personal qualities vs. personal sacrifice for the group) ANOVA on feelings of admiration showed that the main effect of admiration was significant, $F(1, 236) = 857.94$, $p < .001$, $\eta^2_p = .78$. Participants felt more admiration in the admiration conditions ($M = 4.57$, $SD = 0.99$) than in the no admiration conditions ($M = 0.61$, $SD = 1.06$). Neither the effect of reason, $F(1, 236) = 0.71$, $p = .402$, $\eta^2_p = .00$, nor the interaction effect, $F(1, 236) = 0.02$, $p = .876$, $\eta^2_p = .00$, were significant.

Main Analyses

We performed a 2 (no admiration vs. admiration) x 2 (personal qualities vs. personal sacrifice for the group) multivariate analysis of variance (MANOVA) on cost of sacrifice,

fusion with the group, and willingness to fight and die.

Cost of sacrifice. The analysis showed a significant main effect of admiration, $F(1, 236) = 94.43$, $p < .001$, $\eta^2_p = .29$ (admiration conditions: $M = 3.64$, $SD = 1.35$; no admiration conditions: $M = 2.32$, $SD = 0.96$). The main effect of reason was also significant, $F(1, 236) = 13.05$, $p < .001$, $\eta^2_p = .05$ (personal sacrifice conditions: $M = 3.09$, $SD = 1.37$; personal qualities conditions: $M = 2.75$, $SD = 1.27$).

Identity Fusion. The analysis showed a significant main effect of admiration, $F(1, 236) = 7.50$, $p = .007$, $\eta^2_p = .03$ (admiration conditions: $M = 2.39$, $SD = 1.28$; no admiration conditions: $M = 2.03$, $SD = 1.32$). The main effect of reason was not significant, $F(1, 236) = 1.51$, $p = .220$, $\eta^2_p = .01$ (personal sacrifice conditions: $M = 2.23$, $SD = 1.35$; personal qualities conditions: $M = 2.15$, $SD = 1.29$).

Willingness to fight and die. Results showed a marginally significant main effect of level of admiration, $F(1, 236) = 3.55$, $p = .061$, $\eta^2_p = .02$ (admiration conditions: $M = 2.04$, $SD = 0.89$; no admiration conditions: $M = 1.87$, $SD = 0.80$).The main effect of reason was not significant, $F(1, 236) = 1.50$, $p = .223$, $\eta^2_p = .01$ (personal sacrifice conditions: $M = 2.00$, $SD = 0.84$; personal qualities conditions: $M = 1.91$, $SD = 0.86$).

Fusion with other groups. Due to technical issues with the download of the dynamic measures, we could only collect and analyse responses from 126 participants. A 2 X 2 factorial MANOVA on fusion with other groups indicated that the main effect of level of admiration was not significant on fusion with family, $F(1, 122) = 0.24$, $p = .624$, $\eta^2_p = .00$, France, $F(1, 122) = 0.13$, $p = .722$, $\eta^2_p = .00$, Portugal, $F(1, 122) = 0.56$, $p = .457$, $\eta^2_p = .01$, or Italy, $F(1, 122) = 0.20$, $p = .658$, $\eta^2_p = .00$. The main effect of reason was not significant on fusion with family, $F(1, 122) = 0.80$, $p = .374$, $\eta^2_p = .01$, France, $F(1, 122) = 1.34$, $p = .288$, $\eta^2_p = .01$, Portugal, $F(1, 122) = 0.39$, $p = .844$, $\eta^2_p = .00$, or Italy, $F(1, 122) = 0.62$, $p = .434$, $\eta^2_p = .01$. Finally, the effect of the interaction was not significant on fusion with family, $F(1, 122) = 0.03$,

p = .868, η^2_p = .00, France, F(1,122) = 0.14, p = .708, η^2_p = .00, Portugal, F(1, 122) = 0.28, p = .600, η^2_p = .00, or Italy, F(1, 122) = 0.16, p = .693, η^2_p = .00.

Study 6 (N = 771): Long-lasting Effects of Admiration on Real Behaviour

Results

Preliminary Analyses

Differences between both waves. Given the attrition rate (49%), we performed a factorial MANOVA with participation in the second wave (coded 0 = *no*; 1= *yes*) as predictor variable, and intensity of sacrifice and fusion with the group as outcomes.

Results showed that effect of participation in the second wave of the study on cost of sacrifice was not significant, F(1, 769) = 2.57, p = .109, η^2_p = .00. Participants who took part in the second wave did not attribute a different cost of sacrifice to the person they thought about (M = 3.16, SD = 1.25) than participants who did not take part in the second wave (M = 3.31, SD = 1.28). The effect of participation in the second wave on fusion with the country was significant F(1, 769) = 5.62, p = .018, η^2_p = .01. Participants who took part in the second wave were less fused (M = 2.28, SD = 1.05) than participants who did not take part in the second wave (M = 2.47, SD = 1.23).

Feelings of admiration. A unifactorial ANOVA on feelings of admiration showed that the main effect of our manipulation was significant, F (1, 769) = 2339.81, p < .001, η^2_p = .75. Participants felt more admiration in the admiration condition (M = 4.44, SD = 1.30) than in the no admiration condition (M = 0.47, SD = 0.93).

Main Analyses

Second wave

We performed a mixed analysis of variance (ANOVA) with condition as between-subject factor (0 = *no admiration*, 1 = *admiration*) and behaviour (helping vs. aggression) as within-subject factor.

The analyses showed a main significant effect of behaviour, $F(1, 374) = 389.58$, $p < .001$, $\eta^2_p = .51$. Participants exhibited more helping ($M = 5.15$, $SD = 2.99$) than aggressive behaviour ($M = 1.10$, $SD = 1.49$). The main effect of condition was also significant, $F(1, 374) = 36.99$, $p < .001$, $\eta^2_p = .09$, indicating that participants behaved differently in the admiration and no admiration conditions.

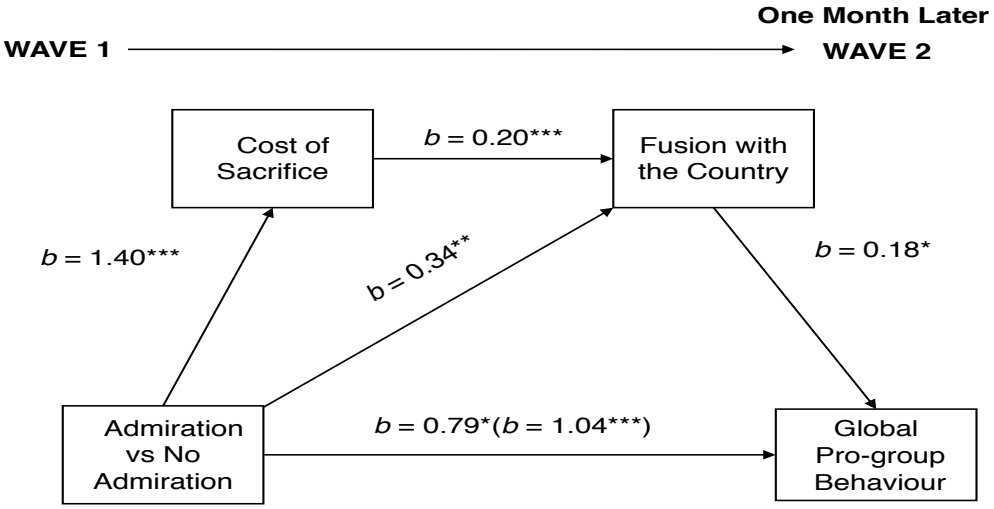
Additional Analysis

We also computed a global measure of pro-group behaviour by multiplying the number of easy tangrams assigned by three, the number of tangrams of medium difficulty assigned by two, and the number of difficult tangrams assigned by one, and by calculating the mean of the multiplied variables, so that higher scores reflect higher pro-group behaviour.

A factorial analysis of variance (ANOVA) on global pro-group behaviour showed that the effect of condition was significant, $F(1,374) = 47.76$, $p < .001$, $\eta^2_p = .11$. Participants' behaviour was more favourable in the admiration condition ($M = 9.33$, $SD = 1.42$) than in the no admiration condition ($M = 8.29$, $SD = 1.50$).

A serial mediation analysis using Hayes'(2018) PROCESS macro Model 6, including condition as predictor, cost of sacrifice as first mediator, identity fusion as second mediator, and global progroup behaviour as outcome variable showed that two of the indirect effects were significant (see Supplementary Figure 1). Condition influenced progroup behaviour via 1) identity fusion alone, $b = 0.06$, 95% CI [0.01, 0.14], and 2) intensity of sacrifice and identity fusion serially, $b = 0.05$, 95% CI [0.01, 0.10]. The indirect effect via cost of sacrifice alone was not significant, $b = 0.14$, 95% CI [-0.07, 0.35].

Figure 13. Linear mediation model in Study 6 (N = 771 vs. 376 for waves 1 and 2)



ANNEX C

Supplementary Materials for Chapter 4: Identity Fusion Predicts Violent Pro-group Behaviour when it is Morally Justifiable

Method Studies 1 and 2

Procedure

We measured willingness to engage in extreme non-violent and violent behaviour with scales clearly differentiated from each other. Both measures include five items, and, in both cases, we asked the participants to imagine that the stability and continuity of the group (i.e., the family in Study 1 and the country in Study 2) are under threat, but we focused on behaviours that entail a strong personal sacrifice without implying violence in the measure of extreme non violent behaviour, whereas we focused on behaviours that entail violence without necessarily implying personal sacrifices in the measure of violent behaviour.

To measure *willingness to engage in non-violent behaviour*, we asked participants to indicate their degree of agreement with the following items: "I would be willing to organize a public demonstration against this threat, although the risk of being attacked by other people was high"; "I would be willing to openly express the idea that the threat should stop, even if I were at risk of being rejected by people important to me"; "I would be willing to quit my job and lose my source of income to organize signature campaigns against this threat"; "I would be willing to give up my freedom for participating in mobilizations and public demonstrations against this threat"; and "I would be willing to die for organizing strikes and demonstrations against this threat". To measure *willingness to engage in violent behaviour*, we asked participants to report their level of agreement with the next items: "If it was useful to end this threat…, I would be willing to attack people"; "…, I would be willing to behave in an aggressive way"; "…, I would be willing to harm people"; …, I would be willing to use violence against other people"; and "…, I would be willing to kill."

To identify the common factors that explain the structure and order among the measured variables, we conducted exploratory factor analyses with oblique rotation. Results of the analyses for the items corresponding to both scales in the two studies are described below:

Study 1

The Kaiser-Meyer-Olkin measure confirmed the sampling adequacy for the analysis, KMO = .87 (Hutcheson & Sofroniou, 1999), and all KMO values for individual items were greater than .74, which is above the conventionally adopted limit of .50 (Field, 2013). An initial analysis was conducted to obtain eigenvalues for each factor in the data, resulting in two factors with eigenvalues over Kaiser's criterion of 1 that conjointly explained 71.83 % of the variance. Table 11 shows the factor loadings after rotation. The clustering of the items suggests that, as expected, the first factor represents willingness to engage in extreme non-violent behaviour and the second factor represents willingness to engage in violent behaviour (see Table 11).

Study 2

The Kaiser-Meyer-Olkin measure showed that the sample was adequate for the analysis, KMO = .85 (Hutcheson & Sofroniou, 1999), and all KMO values for individual items were greater than .76, which is over the acceptable limit of .50 (Field, 2013). An initial analysis was conducted to obtain eigenvalues for each factor in the data, resulting in two factors with eigenvalues over Kaiser's criterion of 1 that conjointly explained 71.94 % of the variance. Table 11 shows the factor loadings after rotation. The clustering of the items indicates that the first factor corresponds to willingness to engage in extreme non-violent behaviour and the second factor corresponds to willingness to engage in violent behaviour (see Table 11).

Table 11. Summary of the exploratory factor analysis for the scales of willingness to engage in extreme non-violent and violent behaviour (Studies 1 and 2)

Rotated Factor Loading

	Study 1. Fusion with Family		**Study 2. Fusion with Country**	
	Non-violent Behaviour	Violent Behaviour	Non-violent Behaviour	Violent Behaviour
Item				
I would be willing to organize a demonstration against this threat, even if the risk of being attacked by other people was high	-.02	**.78**	.02	**.72**
I would be willing to openly express the idea that the threat should stop, even if I were at risk of being rejected by people important to me	-.16	**.67**	-.08	**.66**
I would be willing to quit my job and lose my source of income to organize signature campaigns against this threat	-.01	**.82**	-.06	**.86**
I would be willing to give up my freedom for participating in mobilizations and public demonstrations against this threat	.15	**.78**	.04	**.84**
I would be willing to die for organizing strikes and demonstrations against this threat	.22	**.68**	.37	**.56**
If it was useful to end this threat, I would be willing to attack people	**.92**	.02	**.93**	-.01
If it was useful to end this threat, I would be willing to behave in an aggressive way	**.89**	.01	**.87**	.08
If it was useful to end this threat, I would be willing to harm people	**.97**	-.01	**.97**	-.02
If it was useful to end this threat, I would be willing to use violence against other people	**.96**	.00	**.95**	-.00
If it was useful to end this threat, I would be willing to kill	**.84**	.00	**.86**	-.06
Eigenvalues	4.85	2.33	5.01	2.19
% of Variance	48.55	4.54	50.07	21.87

Note: Factor loadings over .40 appear in bold

References

Field, A. (2013). *Discovering statistics using IBMS SPSS Statistics: And sex and drugs and rock 'n' roll* (4th ed.). Sage.

Hutcheson, G., & Sofroniu, N. (1999). *The multivariate social scientist.* Sage.

ANNEX D

Supplementary Materials for Chapter 5: Strongly Fused Individuals Feel Viscerally Responsible to Self-Sacrifice

Method Studies 1a, 1b, and 2

Sensitivity Power Analyses (Studies 1a and 1b)

We conducted post-hoc sensitivity power analyses with the online tool developed by Schoemann et al. (2017) to determine the statistical power reached by both studies. Considering our sample sizes (N = 200 and N = 182) and the sizes of the effects, we found that Study 1a had 100% power and Study 1b had 99% power (see Figures S2 and S3).

Figure 14. Post-hoc sensitivity power analysis (Study 1a)

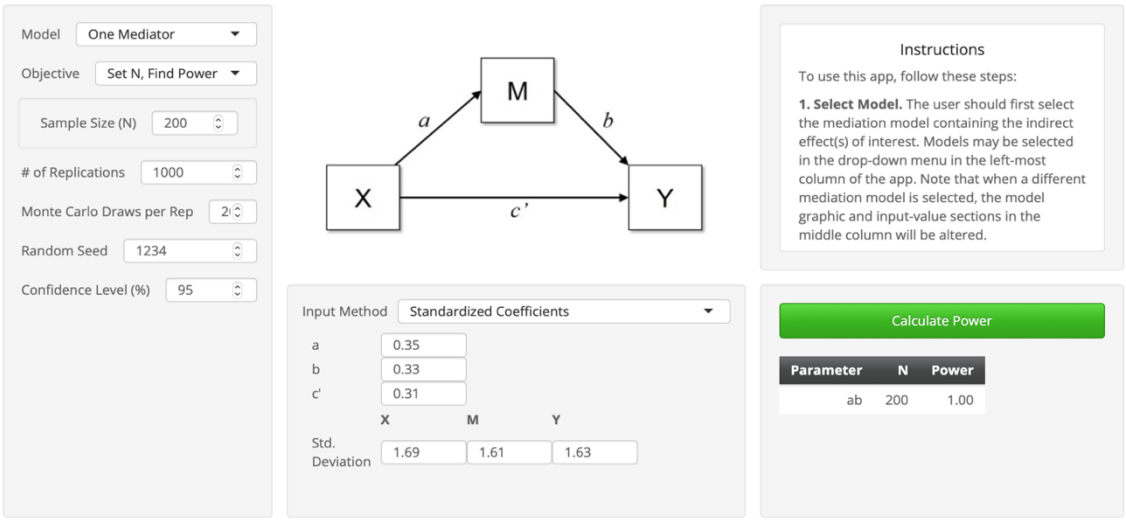
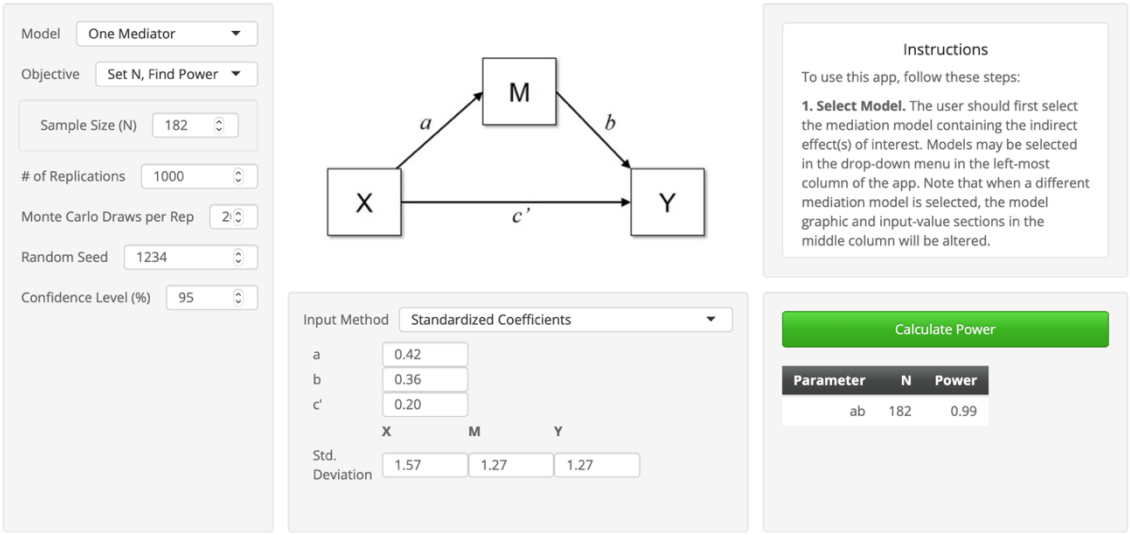


Figure 15. Post-hoc sensitivity power analysis (Study 1b)



Exploratory Factor Analyses

To identify the common factors that explain the structure and order among the measured variables, we conducted exploratory factor analyses on all the items with oblique rotation (direct oblimin). Results of the analyses for the items corresponding to the three scales in all the studies are described below.

Study 1a

The Kaiser-Meyer-Olkin measure confirmed the sampling adequacy for the analysis, KMO = .84 (Hutcheson & Sofroniou, 1999), and all KMO values for individual items were greater than .78, which is over the acceptable limit of .50 (Field, 2013). An initial analysis was conducted to obtain eigenvalues for each factor in the data, resulting in three factors with eigenvalues over Kaiser's criterion of 1 that conjointly explained 71.85 % of the variance. Table 12 shows the factor loadings after rotation. The clustering of the items suggests that, as expected, the first factor represents identity fusion with the partner, the second factor represents feelings of visceral responsibility, and the third factor represents willingness to fight and die for him or her.

Study 1b

The Kaiser-Meyer-Olkin measure showed that the sample was adequate for the analysis, KMO = .86 (Hutcheson & Sofroniou, 1999), and all KMO values for individual items were greater than .77, which is over the limit of .50 (Field, 2013). An initial analysis to obtain eigenvalues for each factor in the data resulted in three factors with eigenvalues over Kaiser's criterion of 1 that conjointly explained 73.24 % of the variance. Table 13 presents the factor loadings after rotation. The clustering of the items indicates that the first factor corresponds to identity fusion with the country, the second factor corresponds to feelings of visceral responsibility, and the third factor corresponds to willingness to fight and die.

Study 2

The Kaiser-Meyer-Olkin measure showed that the sample was adequate for the analysis, KMO = .89 (Hutcheson & Sofroniou, 1999), and all KMO values for individual items were greater than .80, which is over the limit of .50 (Field, 2013). An initial analysis to obtain eigenvalues for each factor in the data resulted in three factors with eigenvalues over Kaiser's criterion of 1 that conjointly explained 74.06 % of the variance. Table 14 shows the factor loadings after rotation. The clustering of the items indicates that the first factor corresponds to identity fusion with the country, the second factor corresponds to feelings of visceral responsibility, and the third factor corresponds to willingness to fight and die.

Confirmatory Factor Analyses

To further assess whether identity fusion, feelings of visceral responsibility and willingness to fight and die should be treated as three separate factors, we conducted confirmatory factor analyses using jamovi (Epskamp, 2017; R Core Team, 2020; Roosel et al., 2018; The jamovi project, 2021). We let the items to load only on the components they were expected to load, and no item errors were permitted to correlate. We compared the fit of our hypothesized three-factor model to that of (1) a two-factor model combining the visceral responsibility and fight and die items; (2) a two-factor model combining the identity fusion and visceral responsibility items; and (3) a single-factor model combining all the items. Results of the analyses are presented below.

Study 1a

As a first step, we modeled the expected three-factor solution with the three fusion items loading on the first factor, the five visceral responsibility items loading onto the second factor, and the three fight and die items loading onto the third factor (for factor loadings and factor intercorrelations see Figure 16). In following steps, we modeled the two- and single-

factor solutions. Results indicated that the three-factor model fitted the data well and has a better adjustment than the two-factor model in which the visceral responsibility and fight and die items loaded onto the same factor ($\Delta\chi^2 = 114$, $\Delta df = 2$, $p < .001$); the two-factor model in which the fusion and visceral responsibility items loaded onto the same factor ($\Delta\chi^2 = 218$, $\Delta df = 2$, $p < .001$); and the single-factor model ($\Delta\chi^2 = 309$, $\Delta df = 3$, $p < .001$). Thus, the three-factor model seems the most appropriate representation of our data among the compared models (See Table 15).

Study 1b

First, we modeled the expected three-factor solution with the three fusion items loading on the first factor, the five visceral responsibility items loading onto the second factor, and the three fight and die items loading onto the third factor (for factor loadings and factor intercorrelations see Figure 17. Next, we modeled the two- and single-factor solutions. Results indicated that the three-factor model fitted the data adequately and has a better adjustment than the two-factor model in which the visceral responsibility and fight and die items loaded onto the same factor ($\Delta\chi^2 = 66.2$, $\Delta df = 2$, $p < .001$); the two-factor model in which the fusion and visceral responsibility items loaded onto the same factor ($\Delta\chi^2 = 210.2$, $\Delta df = 2$, $p < .001$); and the single-factor model ($\Delta\chi^2 = 267.2$, $\Delta df = 3$, $p < .001$). Therefore, the three-factor model seems the most appropriate representation of our data (See Table 16).

Study 2

First, we modeled the expected three-factor solution with the three fusion items loading on the first factor, the five visceral responsibility items loading onto the second factor, and the three fight and die items loading onto the third factor (for factor loadings and factor intercorrelations see Figure 18). Next, we modeled the two- and single-factor solutions. Results indicated that the three-factor model fitted the data relatively well, and has a better adjustment

than the two-factor model in which the visceral responsibility and fight and die items loaded onto the same factor ($\Delta\chi^2 = 118$, $\Delta df = 2$, $p < .001$); the two-factor model in which the fusion and visceral responsibility items loaded onto the same factor ($\Delta\chi^2 = 214$, $\Delta df = 2$, $p < .001$); and the single-factor model ($\Delta\chi^2 = 327$, $\Delta df = 3$, $p < .001$). Thus, the three-factor model seems the most appropriate representation of our data among the compared models (See Table 17).

Table 12. Summary of the exploratory factor analysis for the scales of identity fusion with the romantic partner, feelings of visceral responsibility, and willingness to fight and die for him or her (Study 1a)

Rotated Factor Loadings

	Fusion with the Partner	Visceral Responsibility	Fight and Die
My partner and me are one	**.789**	.028	-.050
I feel immersed in my partner	**.908**	.033	-.014
I feel a strong emotional bond with my partner	**.629**	-.009	.118
I feel impulsively obliged to promote the wellbeing of my partner	-.102	**.927**	-.094
I feel an unreflective duty to take care of my partner	.063	**.840**	-.029
I feel a duty, based on my deepest feelings, to further the interests of my partner	.142	**.728**	.029
I feel an unreflective duty to try to stop something bad from happening to my partner	-.010	**.736**	.113
I feel almost irrationally obliged to act on my partner's whishes	.016	**.686**	.075
I would fight someone who physically threatened my partner	.077	-.039	**.697**
Hurting other people is acceptable if it means protecting my partner	-.112	.084	**.772**
I would sacrifice my life if it saved the life of my partner	.082	.001	**.650**
Eigenvalues	1.84	4.87	1.19
% of Variance	16.75	44.31	10.80

Table 13. Summary of the exploratory factor analysis for the scales of identity fusion with the country, feelings of visceral responsibility, and willingness to fight and die for it (Study 1b)

Rotated Factor Loadings

	Fusion with the Country	Visceral Responsibility	Fight and Die
My country and me are one	**.724**	.099	.051
I feel immersed in my country	**.868**	-.064	-.047
I feel a strong emotional bond with my country	**.818**	.055	.070
I feel impulsively obliged to promote the wellbeing of my country	.028	**.870**	-.093
I feel an unreflective duty to take care of my country	-.067	**.939**	-.038
I feel a duty, based on my deepest feelings, to further the interests of my country	.221	**.681**	.022
I feel an unreflective duty to try to stop something bad from happening to my country	-.073	**.807**	.153
I feel almost irrationally obliged to act on my country's whishes	.054	**.692**	.092
I would fight someone who physically threatened my country	.045	.055	**.699**
Hurting other people is acceptable if it means protecting my country	-.070	.092	**.613**
I would sacrifice my life if it saved the life of another member of my country	.058	-.084	**.632**
Eigenvalues	1.69	5.08	1.30
% of Variance	15.32	46.14	11.77

Table 14. Summary of the exploratory factor analysis for the scales of identity fusion with the country, feelings of visceral responsibility, and willingness to fight and die for it (Study 2)

Rotated Factor Loadings

	Fusion with the Partner	Visceral Responsibility	Fight and Die
My country and me are one	**.572**	.172	.107
I feel immersed in my country	**.944**	-.061	-.074
I feel a strong emotional bond with my country	**.794**	.022	.042
I feel impulsively obliged to promote the wellbeing of my country	.008	**.851**	-.005
I feel an unreflective duty to take care of my country	-.062	**1.02**	-.086
I feel a duty, based on my deepest feelings, to further the interests of my country	.216	**.574**	.166
I feel an unreflective duty to try to stop something bad from happening to my country	-.022	**.836**	.102
I feel almost irrationally obliged to act on my country's whishes	.111	**.624**	.038
I would fight someone who physically threatened my country	.065	.003	**.786**
Hurting other people is acceptable if it means protecting my country	-.030	-.072	**.861**
I would sacrifice my life if it saved the life of another member of my country	.139	-.011	**.420**
Eigenvalues	1.35	5.75	1.04
% of Variance	12.30	52.31	9.45

Table 15. Fit-Indices for model comparison (Study 1a)

Fit Statistic	Three factor model	Two-factor model (responsibility+fight-die)	Two-factor model (fusion+responsibility)	One-factor model
χ^2(df)	102***(41)	216***(43)	320***(43)	411***(44)
CFI	0.95	0.84	0. 75	0.67
TLI	0.93	0.80	0.68	0.58
SRMR	0.05	0.10	0.12	0.13
RMSEA (95% CI)	0.08 (0.06-0.11)	0.14 (0.12-0.16)	0.18 (0.16-0.20)	0.20 (0.19-0.22)

*** $p < .001$

Table 16. Fit-Indices for model comparison (Study 1b)

Fit Statistic	Three factor model	Two-factor model (responsibility+fight-die)	Two-factor model (fusion+responsibility)	One-factor model
χ^2(df)	90.8***(41)	157***(43)	301***(43)	358***(44)
CFI	0.95	0.89	0.76	0.71
TLI	0.94	0.86	0.69	0.64
SRMR	0.05	0.08	0.11	0.12
RMSEA (95% CI)	0.08 (0.06-0.10)	0.12 (0.10-0.14)	0.18 (0.16-0.20)	0.20 (0.18-0.22)

*** $p < .001$

Table 17. Fit-Indices for model comparison (Study 2)

Fit Statistic	Three factor model	Two-factor model (responsibility+fight-die)	Two-factor model (fusion+responsibility)	One-factor model
χ^2(df)	146***(41)	264***(43)	360***(43)	473***(44)
CFI	0.95	0.89	0. 84	0.79
TLI	0.93	0.86	0.80	0.74
SRMR	0.05	0.07	0.08	0.09
RMSEA (95% CI)	0.09 (0.07-0.11)	0.13 (0.11-0.15)	0.16 (0.14-0.17)	0.18 (0.16-0.19)

*** $p < .001$

Figure 16. Confirmatory factor analysis of the three-factor model (Study 1a)

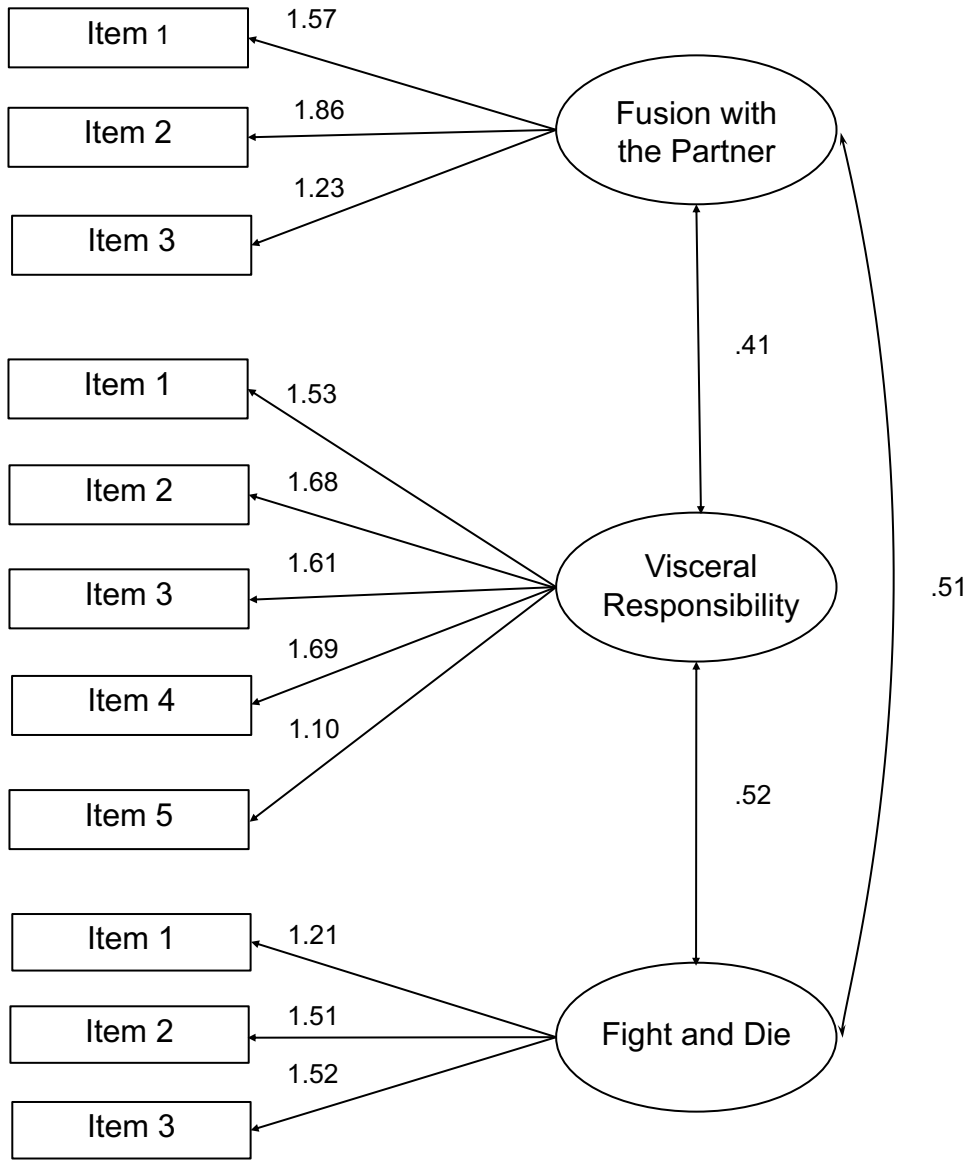


Figure 17. Confirmatory factor analysis of the three-factor model (Study 1b)

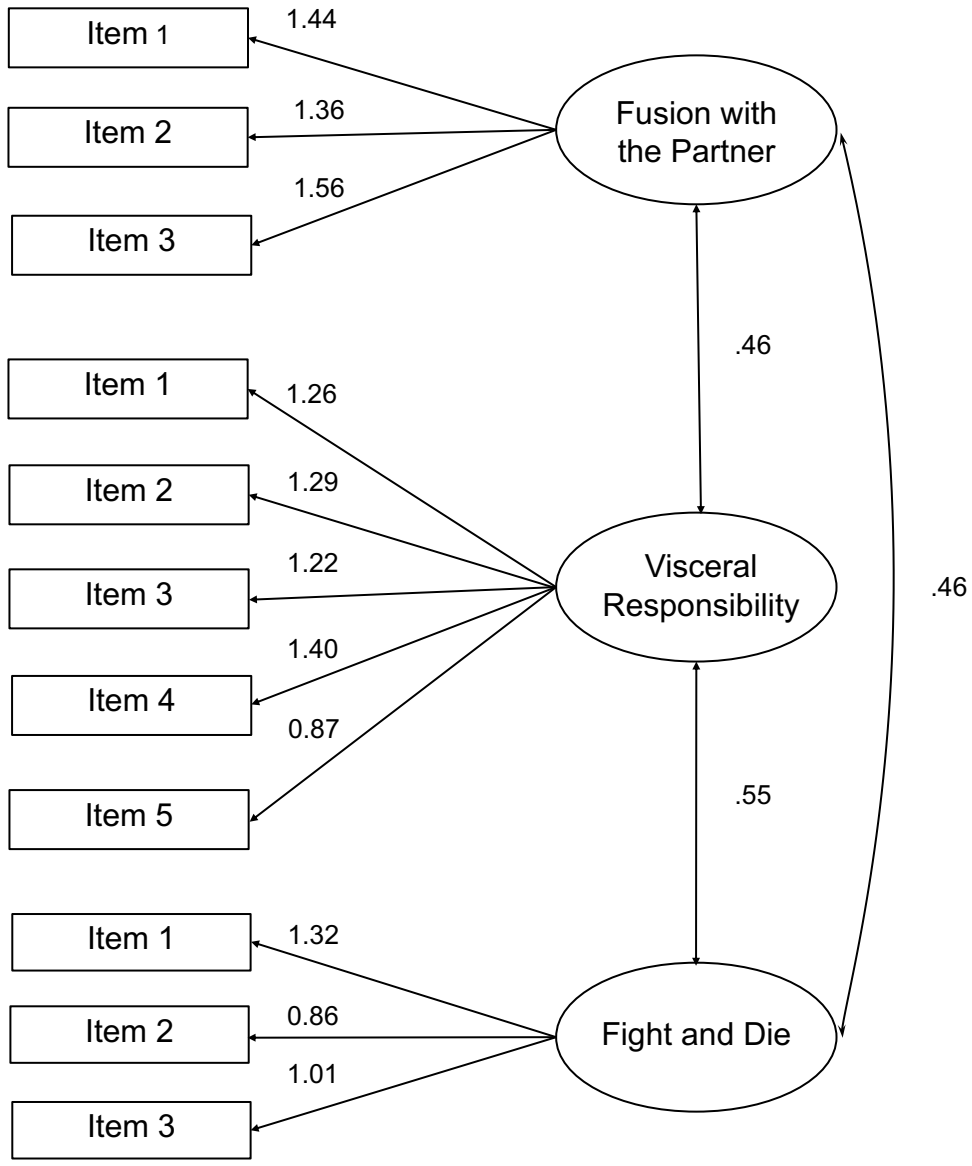
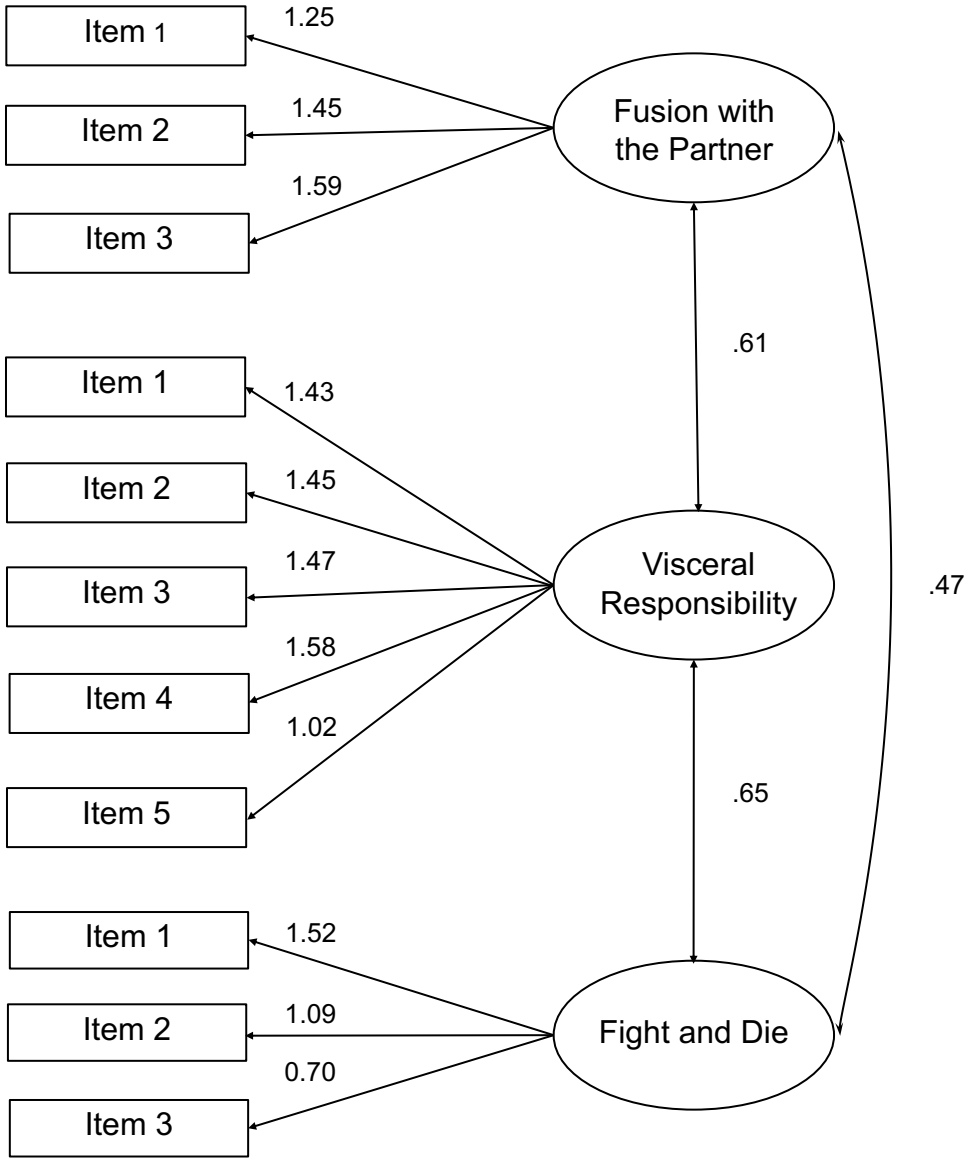


Figure 18. Confirmatory factor analysis of the three-factor model (Study 2)



References

Epskamp, S. (2017). *semPlot: Path diagrams and visual analysis of various SEM packages' output.* [R Package]. https://cran.r-project.org/package=semPlot.

Field, A. (2013). *Discovering statistics using IBM SPSS Statisctics* (4th Ed.). Sage.

Hutcheson, G., & Sofroniu, N. (1999). *The multivariate social scientist.* Sage.

R Core Team (2020). *R: A Language and environment for statistical computing.* (Version 4.0) [Computer Software]. https://cran.r-project.org

Rosseel, Y., et al. (2018). *Lavaan: Latent Variable Analysis.* [R Package]. https://cran-r-project.org/package=lavaan

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science, 8*(4), 379-386. https://doi.org/10.1177/1948550617715068

The jamovi project (2021). jamovi. (Version 1.6) [Computer Software]. https://jamovi.org


UNED
MOBILIBVS
MOBILIOR
SAPIENTIA
OMNIBVS