Evolution and Human Behavior 43 (2022) 408–417

Contents lists available at ScienceDirect

Evolution and Human Behavior

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

Does hazing actually increase group solidarity? Re-examining a classic theory with a modern fraternity☆

Aldo Cimino [a,*], Benjamin J. Thomas [b]

[a] Department of Anthropology, Kent State University, OH, United States of America
[b] Center for Leadership and Ethics, McCombs School of Business, University of Texas at Austin, TX, United States of America

ARTICLE INFO

Keywords:
Hazing
Newcomers
Rites of passage
Fraternities

ABSTRACT

Anthropologists and other social scientists have long suggested that severe initiations (hazing) increase group solidarity. Because hazing groups tend to be highly secretive, direct and on-site tests of this hypothesis in the real world are nearly non-existent. Using an American social fraternity, we report a longitudinal test of the relationship between hazing severity and group solidarity. We tracked six sets of fraternity inductees as they underwent the fraternity's months-long induction process. Our results provide little support for common models of solidarity and suggest that hazing may not be the social glue it has long been assumed to be.

1. Introduction

Hazing—the abuse of new or prospective group members—has fascinated researchers for over a century, leading to a variety of different theories and speculations about its purpose and effects on participants (e.g., Durkheim, 1912; Tiger, 1984; Van Gennep, 1909; Whitehouse, 1996). Hazing ordeals can vary widely, and may include intoxication, privation, humiliation, servile labor, beatings, and brandings. Hazing is notable for its wide cross-cultural distribution, historical depth, and modern persistence despite concerted efforts at moralization and suppression (e.g., Allan & Madden, 2012; Barber, 2012; Cimino, 2020). There are three common themes in the explanations given for hazing's genesis or persistence: 1. hazing creates group solidarity, 2. hazing is an expression of dominance, and 3. hazing allows for the selection of committed group members (for a review, see Cimino, 2011). These are the "macro theories" of hazing: The three broad ideas that continually recur throughout the social sciences. Nearly all claims about hazing's function (actual or perceived) are categorizable under at least one of these three theories, though with varying specifics and scope (e.g., street gang inductions, adolescent initiations). While the macro theories make for a quick thematic summation of decades of prior work, we emphasize that researchers tend to have highly nuanced perspectives on hazing in any given group, and the macro theories are not an exhaustive description of all relevant theorizing (e.g., Cimino, 2016; Grimes, 2000). Indeed, hazing initiations are widely recognized as complex and multivocal, and may be enmeshed in cultural processes relevant to maturation and gender (e.g., Herdt, 1998), warfare (e.g., Ember & Ember, 2010), and religion (e.g., Gill, 1996). Nonetheless, in order to systematically build a scientific understanding of hazing practices, we must begin to isolate specific claims and subject them to formal testing. And by far the most prominent, generalizable, and enduring claim about hazing is that it increases group solidarity. In this article, we report a quantitative test of this prediction with an American social fraternity. We longitudinally surveyed six sets of fraternity inductees (i.e., "pledge classes") as they underwent the fraternity's months-long induction process. We gathered data on hazing severity and multiple measures of solidarity, allowing us to test for any relationships therewith. To the best of our knowledge, this study represents the most comprehensive test of hazing's relationship with solidarity yet performed.

In addition to providing a key empirical test of an important and enduring claim, our findings can contribute to the refinement of evolutionary theories of hazing and hazing-inclusive phenomena (e.g., costly or dysphoric rituals). For example, if hazing substantially increases group solidarity, this would be consistent with the idea that hazing (among many similar dysphoric practices) may have been culturally group selected (Whitehouse & Lanman, 2014). But if hazing is poor or notably inefficient at creating feelings of group solidarity, it might suggest that the evolution of hazing behaviors (cultural or

☆ We thank Steven J. C. Gaulin and Kristin Yeager for helpful suggestions and Deanna Nathan for research assistance. Special gratitute is extended to the members of Beta, in particular the PI's primary contact, "Odysseus".

* Corresponding author at: Department of Anthropology, Kent State University, 750 Hilltop Drive, 226 Lowry Hall, Kent, OH 44242, United States of America.
E-mail address: acimino1@kent.edu (A. Cimino).

https://doi.org/10.1016/j.evolhumbehav.2022.07.001
Received 7 February 2022; Received in revised form 14 May 2022; Accepted 15 July 2022
Available online 29 July 2022

genetic) was for reasons largely unrelated to solidarity (e.g., Cimino, Toyokawa, Komatsu, Thomson, & Gaulin, 2019), or even that hazing behaviors may be a kind of byproduct.

1.1. Defining hazing

We have glossed hazing as "the abuse of new or prospective group members", but this is not a formal definition of the phenomenon. For the purposes of our study and literature review, we employed the "strict" definition of hazing (Cimino, 2007, 2017):

> Hazing is non-accidental, costly aspects of group induction activities that: a) do not appear to be group-relevant assessments/preparations, or b) appear excessive in their application. Group induction activities are those tasks formally or informally required to obtain membership or participatory legitimacy for new or prospective members. (Cimino, 2017, p. 144)

For example, requiring inductees to perform calisthenics is a group-relevant assessment or preparation for athletic teams, but not for social fraternities. Importantly, the strict definition exists to circumscribe the practices under investigation, not to presuppose their ultimate nature. That is, we are not claiming that hazing *necessarily* lacks group relevance, only that it gives such an appearance. It may well turn out that hazing is in some way cryptically group relevant (e.g., Keating et al., 2005; Precourt, 1975). But if so, this must be discovered rather than assumed. For more on the definition of hazing, see Cimino (2017) and Thomas, Cimino, and Meglich (2021).

1.2. The solidarity macro theory and the American fraternity

The solidarity macro theory is the idea that hazing increases social harmony, cohesion, coordination, or the like. Given such a description, one might wonder whether the commitment and dominance macro theories also entail increases in group solidarity. For example, selecting out less-committed members might indirectly increase social harmony. But in order to be considered part of the solidarity macro theory as we have defined it, the hypothesized manner in which hazing generates solidarity must be logically separable from the other two macro theories. That is, we are investigating the proposition that hazing directly transforms hazees into intrinsically dedicated members.

For the solidarity macro theory, hazing's effects are usually stated or implied to be driven by psychological changes in hazees, whereby hazing causes increases in solidarity-relevant cognitions (e.g., liking, trust, dedication, identification). The solidarity macro theory encompasses different claims about the targets of these hazee cognitions. Specifically, solidarity might be increased toward (a) fellow hazees, (b) hazers, or (c) the group as a whole.[1] When applied to American social fraternities, these categories are as follows:

Fellow hazees ("pledges"): Pledges are formally recognized fraternity inductees. That is, individuals who have been pre-approved by incumbent members to collectively take part in the group's induction process. Once they have completed the process to the satisfaction of incumbents, they become "actives".

Hazers ("actives"): Actives are incumbent members of the fraternity that have not yet graduated (after which they become fraternity alumni). Fraternity hazing is conducted by active members, though alumni may sometimes participate.

The group as a whole ("chapter"): National fraternities are broken up into school-specific chapters. For example, the first American fraternity, Phi Beta Kappa, has 290 chapters.

Given these definitions, we tested a straightforward prediction made by the solidarity macro theory:

The severity of hazing experienced by fraternity pledges will be positively associated with their feelings of solidarity for fellow pledges, actives, the chapter, or some subset of these targets.

Additionally, we suggest that the explanatory power of the solidarity macro theory depends on one or more of these associations being moderate to large in effect size. That is, it makes little sense to claim that hazing processes are lengthy and intense because they tend to *ever-so-slightly* increase feelings of group solidarity. This is especially so given past treatments of severe initiations as transformative bonding experiences (e.g., Tiger, 1984; Turner, 1967; Whiting, Kluckhohn, & Anthony, 1958; Young, 1965).

1.3. Why hazing might increase group solidarity

Prior to reviewing the relevant literature, it is worth briefly describing why hazing might cause an increase in group solidarity. There are at least four reasons that have either prima facie plausibility or some systematic evidence in their favor:

1.3.1. Increased opportunities to build trust

Hazing ordeals can take many forms that could, in principle, allow hazees to increase their trust for one another. For example, the pledges of the pseudonymous fraternity "Alpha" were made to perform calisthenics (Cimino, 2016), but were sometimes allowed to provide assistance to fellow pledges who were tiring. This kind of assistance, in addition to commiseration or other forms of direct or indirect emotional support, may hasten what would otherwise be a slower trust-building process (e.g., Lodewijkx & Syroit, 1997; Schachter, 1959).

1.3.2. Effort justification via cognitive dissonance

Sometimes called the "severity-attraction hypothesis", this is the idea that hazing creates cognitive dissonance in hazees (Aronson & Mills, 1959). Specifically, because hazing ordeals are very unpleasant, the fact that hazees have endured such ordeals is dissonant (i.e., inconsistent) with their recognition that aspects of the hazing group are suboptimal and not worth the effort. Hazees may resolve this dissonance by deciding that they like the group more than they would have otherwise, thus internally justifying their own effort.

1.3.3. Hazer-directed dependence

The dependence hypothesis is the idea that the maltreatment of hazees can create paradoxically positive sentiments toward hazers, perhaps via attachment-related motivational systems (e.g., Keating et al., 2005; Schopler & Bateson, 1962). This would make hazing a process that can create something akin to Stockholm Syndrome in hazees.

1.3.4. Identity fusion

Identify fusion theory (Swann, Jetten, Gómez, Whitehouse, & Bastian, 2012) posits that when group members have shared dysphoric experiences, it can cause them to "fuse" their identity with the group: That is, to treat their own interests as nearly coextensive with the group's interests and thus increase their willingness to sacrifice for the group. The process of dysphoria-based fusion is potentially complex (Reese & Whitehouse, 2021), but is considered at least partly separable from the other mechanisms listed here (e.g., cognitive dissonance). Severe initiations have been characterized as one of many real-world generators of identity fusion (Whitehouse et al., 2017).

The above four processes do not exhaust all the pathways through which hazing might increase group solidarity, and any number of these processes might be simultaneously operative. For example, hazees might experience increased opportunities to build trust among one another *and* maltreatment-derived feelings of dependence on hazers. But no matter

[1] Because groups that haze change their membership composition over time, members may have a mental model of "the group" as an enduring entity that is separable from its visible membership at any given time point. This is particularly relevant to fraternity chapters, as active members regularly graduate and are replaced with new active members.

which processes are operative, they all predict measurable increases in one or more solidarity variables, directed at one or more targets (e.g., trust among hazees, liking of hazers).

Our summaries of the above mechanisms are brief because our study was not designed to differentiate between them. Instead, for our sample, we intended to comprehensively test for any positive association between hazing and group solidarity, regardless of the underlying mechanism(s). Despite common beliefs to the contrary, whether such an association exists is the primary, unanswered scientific question.

1.4. What we know about hazing and solidarity

Below we review quantitative studies that measure either hazing's effect on group solidarity or hazing's observed association with group solidarity. It is already established that there are many qualitative and high-level observations that appear consistent with the idea that hazing creates group solidarity (see cites in Cimino, 2011), but our focus is explicitly on more formalized, individual-level measurement and testing.

1.4.1. Naturalistic surveys

Most of the relevant survey data on hazing has been gathered from US college students or subsets thereof. Much of it is also challenging to interpret with confidence. For example, Owen, Burke, and Vichesky (2008) surveyed members of collegiate student groups, and found that individuals who reported exposure to more hazing practices (and thus likely experienced more severe hazing) were more likely to agree with the statement that hazing "allows probationary members to bond". This association might suggest that severe hazing increases feelings of solidarity, though it might simply suggest that individuals in organizations with more severe hazing are more prone to justify it in terms of solidarity (see McCreary & Schutts, 2019).

Keating et al. (2005) also surveyed members of collegiate student groups about their initiation experiences. Keating et al. aggregated their measures into separate ratings of initiation deviance, fun, and harshness. Both fun and harshness positively predicted the importance of the group to participants (see also Whitehouse et al., 2017, p. 6). With respect to harshness, this is again the kind of association one would expect if hazing were creating solidarity, but it is also the case that groups that haze more severely may tend to be higher-quality groups (i.e., more group benefits, see Cimino, 2013b). Higher-quality groups may also tend to be more important to their members, and thus this association may not be a direct product of hazing. Indeed, Mann, Feddes, Doosje, and Fischer (2016) conducted a conceptually similar survey of Dutch fraternity and sorority members that asked about affiliation (i.e., bonding) rather than personal importance. They found that members' rated recollections of different types of hazing had either negative or non-significant associations with their rated recollections of bonding during their initiation and at present.

Rogers, Rogers, and Anderson (2012) surveyed alumni of a black Greek letter organization, some of whom had experienced a hazing induction, and some of whom had not. After applying a number of reasonable controls (e.g., annual income), they found that exposure to hazing was not a significant predictor of post-college fraternity involvement, as measured by paying monthly dues to their national organization. While maintaining one's membership is a legitimate measure of dedication, Rogers et al. acknowledged that any enduring solidarity experienced by members might be more strongly felt toward fellow inductees and their chapter, not their national organization (see Parks & Brown, 2005).

Van Raalte, Cornelius, Linder, and Brewer (2007) surveyed US student athletes, finding either negative or non-significant relationships between their extent of self-reported hazing and sub-scales of a team cohesion measure (the Group Environment Questionnaire). We note Van Raalte et al.'s findings only because they are an expected citation for our research question. It is not widely understood that Van Raalte et al. used an idiosyncratic operationalization of hazing that is not compatible with most hazing definitions, including our own (see Cimino, 2017, pp. 142–143).

There is also some initial research on workplace hazing in the United States. Most relevant is Mawritz, Capitano, Greenbaum, Bonner, and Kim (2020)'s recently developed scale to assess experiences of hazing in occupational workgroups from the perspective of hazees. They found that each of their measure's sub-scales of hazing exposure, when assessed simultaneously, either non-significantly or negatively predicted organizational commitment (but see caveats in Thomas et al., 2021).

In sum, the relevant survey studies, while few in number, do not collectively suggest a clear and obvious positive association between hazing and group solidarity. However, in some cases, these studies may have been limited by simple measures, their retrospective nature (e.g., errors in recall), or insufficient control variables. In order to reliably detect the putative effect of hazing, we may need to start at the point of its ostensible instantiation.

1.4.2. Experiments and quasi-experiments

One way to start at the point of ostensible instantiation is to examine laboratory experiments that attempt to simulate hazing. The first such experiment was famously conducted by Aronson and Mills (1959). Aronson and Mills proposed the aforementioned cognitive dissonance theory of hazing. They tested it with an all-female group of US undergraduates joining an intentionally boring, sham discussion group. Aronson and Mills found that participants who experienced a "severe" initiation (i.e., reading embarrassing words/erotica out loud) reported liking the group more than those who experienced a mild initiation, or no initiation at all. A few early attempts to replicate Aronson and Mills (1959) succeeded, while later attempts wholly or partially failed (see reviews and potential explanations in Enge, 1993; Lodewijkx & Syroit, 1997; but see also Keating et al., 2005). We do not have the space to recapitulate all the potential issues. However, we do want to highlight that, from an anthropological perspective, the Aronson and Mills paradigm was best designed to test the theory of cognitive dissonance, not to closely match the actual severity and context of many real-world hazing processes (cf. Keating et al., 2005). As Cimino (2013a) put it:

> Compare [laboratory hazing ordeals] to the *months* of hazing experienced by pledges of [pseudonymous fraternity "Alpha"]: the constant barrage of intimidation and yelling; the vomit, sweat, and bloody knuckles from brutal calisthenics and nauseating food; the fear, exhaustion, thirst, pain, and tedium. In juxtaposition, it is difficult to see how experimentally induced "severe" initiations can be treated as unproblematic microcosms of genuinely severe hazing. While there are likely continuities in the psychological impacts of both minimal and maximal hazing, it seems equally likely that there are dramatic discontinuities as well, akin to those between stubbing a toe and losing a leg. (p. 137)

Though it is understandable why "severe" laboratory initiations are actually relatively trivial, ethical boundaries are not a license to disregard external validity when interpreting findings. This issue persists, with the most recent attempt to replicate Aronson and Mills (1959) using a severe initiation that consisted of a rudely assessed five-minute math quiz (Kamau, 2013; see also Keating et al., 2005, Study 2 and 3). It is also unclear that the experience of joining strange and ephemeral laboratory groups is a good psychological proxy for the experience of joining the kind of serious and enduring coalitions that tend to engage in non-trivial hazing in the real world.

In sum, hazing simulation experiments suggest that brief and relatively mild instances of hazing might *sometimes* have positive effects on solidarity in odd and artificial circumstances. While this is interesting, it has not been convincingly shown to scale up to the kind of real-world hazing processes that are of primary interest. Consider Lodewijkx and Syroit (1997), whose study was the first attempt to directly and longitudinally test the inferences from the Aronson and Mills paradigm in the

real world. Lodewijkx and Syroit performed a study using inductees to two Dutch student groups: One sorority with a severe initiation and one mixed-sex, historically religious group with a milder initiation. (In the Netherlands, student groups can have lengthy and unpleasant inductions, much like American Greek letter societies. These can involve common hazing ordeals like privation, sleep deprivation, and humiliation.) Lodewijkx and Syroit took multiple survey measurements that included operationalizations of initiation severity as well as perceptions of liking among inductees. The upshot of their analyses was that, for both the severe and mild student groups, perceived initiation severity actually predicted *less* liking of fellow inductees. Lodewijkx and Syroit (2001) then reanalyzed their data to more directly compare the two groups, finding that inductees to the severe group did not report more liking of fellow inductees than the mild group.

1.4.3. Summary

Despite there being a number of plausible reasons why hazing might increase solidarity, and despite a number of relevant studies, the available quantitative evidence that hazing actually does so is often negative or equivocal. Some of the problems in the relevant literature are greater than others, in particular (a) the difficulty of interpreting extant retrospective surveys, (b) the lack of external validity in laboratory experiments, and (c) the sometimes limited and simplistic manner in which solidarity has been measured thus far. Our study was designed to mitigate these issues by (a) measuring the ostensible effect of hazing during the hazing process, (b) examining a real-world group rather than a laboratory group, and (c) measuring solidarity more comprehensively than any hazing study to date.

2. Methods

All participants were pledges to the pseudonymous fraternity chapter "Beta", located at an anonymous university in the United States. In order to conduct the study, the principal investigator established a rapport with an active member of Beta. Through the PI's relationship with this active member and subsequent conversations with other Beta actives, the chapter formally agreed to participate in a longitudinal survey study of their induction process. The study period covered their approximately ten-week induction, with pledges filling out the same survey at five time points. Each anonymous survey measured pledges' self-reported ratings of the harshness and fun of their induction and self-reported ratings of solidarity (see Measures for details). This process was repeated for six different Beta pledge classes between January 2012 and October 2014 (total $N = 126$, Table 1).

For each pledge class, the study proceeded as follows:

The PI would communicate with a representative of Beta. The representative and the PI would come to an agreement on the specific date of the initial four survey time points. The 1st time point was always during the first week of pledging. The 5th time point was always during the last week of pledging, and because this could vary, the 5th time point was agreed upon once the induction was closer to completion. Separation between time points was intended to be around two weeks, but maximal flexibility was given to Beta to determine specific dates that made sense in the context of the induction. Because of this fact, time points could vary considerably. For example, pledge classes E and F reached the 4th time point near the end of a normal academic school year, and thus their induction processes (and the 5th time point) were put on hold until Fall.

For time point 1 of each study period, the PI would visit the fraternity house and pay the chapter $250 (for pledge class A) or $265 (for all subsequent pledge classes).[2] The PI would then introduce himself to the pledge class, and describe how the study would proceed. The PI noted that pledges could choose whether to participate, that there were no right or wrong answers, and that individual survey responses would be anonymous and unknown to the actives. Pledges were offered $5–$10 for their participation, depending on available funding during the study period (see next section). The PI would then distribute the 1st time point survey to the pledges, wait until all participants had completed it, and then collect the surveys and pay the participants. The PI would provide Beta with a large plastic tub that contained the surveys for time points 2, 3, and 4. (Time point 5 was done in-person, like time point 1; see below.) The tub also contained stamped envelopes with the PI's mailing address for the contained surveys. Printed on the tub was a reminder of the agreed-upon survey dates, and contact information for both the researcher and local Institutional Review Board.

For time points 2, 3, and 4, the PI was not present. Instead, Beta would direct the pledges to somewhere of their choosing to fill out the surveys. Beta agreed to ensure that actives were never present while the pledges filled out their surveys, and additionally agreed that they would never ask pledges about their answers. In order to increase pledge confidence in anonymity, two steps were taken. First, pledges did not write their names on their surveys. Instead, pledges wrote their favorite color and mother's birthday. Second, for time points 2, 3, and 4, pledges placed their surveys inside the aforementioned stamped envelopes and took them to a nearby mailbox after completion.

For time point 5, the PI once again visited the fraternity house in-person and gave the chapter a second payment of $250 (for pledge class A) or $265 (for all subsequent pledge classes). As with time point 1, the PI distributed the survey, collected the completed versions, and (for pledge classes A, C, D, and F) paid an additional $5 to participants.

2.1. Qualifications and challenges

Studying real-world groups always presents challenges, and the above process has some qualifications. Here we note both general issues encountered and differences across pledge classes:

(a) Fraternity induction processes appear to be standardized within chapters (e.g., Cimino, 2016; McMinn, 1980). That is, chapters seem to have some agreed-upon ordering of activities or ordeals that must be completed. However, chapters can also make decisions to change around at least some aspects of the induction for some pledge classes. Further, across different pledge classes, Beta may have inconsistently placed survey time points relative to completed activities. As such, the five time points used in this study should only be regarded as roughly equivalent progress markers. That is, it should not be assumed that all pledges at a

Table 1
Response volume/attrition rates across time points.

	Response volume (i.e., count) at time				
Pledge class	1	2	3	4	5
A	15	10	9	9	9
B	15	10	7	7	7
C	41	28	23	21	22
D	8	5	0	5	5
E	20	13	8	9	7
F	27	17	16	15	17
Total	**126**	**83**	**63**	**66**	**67**

Note. Total study attrition was closely matched to induction attrition. That is, participants who did not respond at time point 5 had very likely dropped out of the pledging process. This is inferred because time points 1 and 5 were both done in-person, allowing the PI to observe pledges filling out the surveys. While the PI did not observe any non-participants at these time points, these observations are not faultless: Some pledges may have been absent due to sickness or other issues. It is also possible that at least a few non-participants were missed for other reasons, such as inattention by the PI.

[2] Chapter payments were made to collectively reimburse actives for their time, as the study process required their cooperation and necessarily inconvenienced them.

given time point experienced the exact same induction activities (see also (b) below).

(b) For time points 2, 3, and 4, pledges were instructed to seal their surveys in the provided pre-addressed, postage-paid envelopes and then take them outside to a nearby mailbox. While this instruction appears to have largely been followed, a minority of surveys were sent late, presumably because some participants were not present when the pledges were originally given the surveys for that time point (e.g., due to sickness), or forgot to send their surveys upon completion.

(c) Across pledge classes, all surveys were identical save the surveys for pledge class A, which did not have a "Today's Date" field for pledges to manually fill out. This field was added to all subsequent surveys to get an additional date indicator other than the initially agreed-upon survey dates.

(d) Pledges were instructed not to talk to one another while filling out surveys, but it turns out that getting a room full of pledges to stop chatting with one another is near impossible. As such, surveys should not be regarded as being completed in quiet, isolated contemplation.

(e) As noted above, for each survey, pledges wrote down their mother's birthday and favorite color. This was so pledges in a given pledge class could be matched across time points. But pledges only had a very general sense of the date of their mother's birth, and (to our surprise) their favorite color sometimes changed over time. This meant that many surveys had to be matched via handwriting. The PI worked with a research assistant to match surveys. Those that could not be matched (6) were excluded from analysis. In addition, one survey sent was a duplicate and two surveys were found completed but never sent. These were also excluded. Pledges would occasionally write incomplete or obviously incorrect dates (e.g., writing 2015 when the year was 2014), requiring the actual date to be inferred from context (e.g., the time point number on the survey and its grouped arrival with other surveys). Finally, pledges would sometimes circle two adjacent numbers on a rating scale (these were entered as averages). For our complete dataset and stimuli, see http://www.aldocimino.com/solidarity_data.zip.[3]

(f) This study was personally funded by the PI. As a consequence, available funds varied, and participants in pledge classes B and E were paid $5 rather than $10.

(g) Pledge class D never filled out the survey for time point 3, as the actives simply forgot to distribute it.

2.2. Measures

As a reminder, we are investigating the following prediction:

The severity of hazing experienced by fraternity pledges will be positively associated with their feelings of solidarity for fellow pledges, actives, the chapter, or some subset of these targets.

Below we list our induction and solidarity measures.

2.2.1. Induction measures

2.2.1.1. Induction harshness. Induction Harshness consisted of four items using 7-point rating scales (1 = Not at all; 7 = Very Much), and was designed to be our operationalization of hazing severity. All items consisted of the phrase "How [adjective] is the pledging process?", where the adjectives were harsh, stressful, tough, and unpleasant. Note that it is almost always the case that when fraternity inductions are in any way harsh, it is because they are integrating hazing ordeals, and hazing was known by the PI to be taking place in Beta.

2.2.1.2. Induction fun. Induction Fun was constructed identically to Induction Harshness, except the adjectives were fun, entertaining, enjoyable, and pleasant. It is important to have a measure like Induction Fun because not every fraternity induction activity qualifies as hazing, or is even intended to be an ordeal. There can be genuinely enjoyable interludes in the midst of a fraternity's hazing induction process (e.g., parties), or just activities that are straightforward, group-relevant, and interesting (e.g., learning about the chapter's history).

2.2.2. Pledge class-directed solidarity measures

2.2.2.1. Pledge class liking. Pledge Class Liking consisted of four items using 7-point rating scales (1 = Not at all; 7 = Very Much). Participants indicated how much they liked the members of their pledge class, responding to items like "The members of my pledge class feel like family" and "I like the individual members of my pledge class".

2.2.2.2. Pledge class coordination. Pledge Class Coordination consisted of four items using 7-point rating scales (1 = Not at all; 7 = Very Much). Participants indicated how well their pledge class worked together, responding to items like "Members of my pledge class find it easy to work together towards a common goal" and "My pledge class works efficiently at group tasks".

2.2.2.3. Pledge class communal strength. (adapted from Mills, Clark, Ford, & Johnson, 2004) Pledge Class Communal Strength consisted of 10 items using 10-point rating scales (1 = Not at all; 10 = Extremely). Participants indicated how much dedication they felt to their pledge class, responding to items like "How happy do you feel when doing something that helps the pledge class?" and "How high a priority for you is meeting the needs of the pledge class?"

2.2.2.4. Pledge class other-in-self. (adapted from Aron, Aron, & Smollan, 1992) Pledge Class Other-in-Self consisted of a single item. Participants indicated their closeness with their pledge class by selecting one of seven pictures. Each picture showed two circles representing the participant and the pledge class, with the circles initially non-overlapping (1) and progressing to be almost completely overlapping (7).

2.2.3. Active and chapter-directed solidarity measures

2.2.3.1. Active liking. Active Liking was constructed identically to Pledge Class Liking, except "members of my pledge class" was replaced with "actives" or "the actives" (e.g., "The actives feel like family", "I like the individual actives").

2.2.3.2. Chapter communal strength. (Adapted from Mills et al., 2004) Chapter Communal Strength was constructed identically to Pledge Class Communal Strength, except "pledge class" was replaced with "chapter" (e.g., "How high a priority for you is meeting the needs of the chapter?").

2.2.3.3. Chapter other-in-self. (Adapted from Aron et al., 1992) Chapter Other-in-Self was constructed identically to Pledge Class Other-in-Self, except "pledge class" was replaced with "the chapter".

2.3. Data analysis

To test the interrelationships between survey items, responses to items were processed using a factor analysis of the complete dataset: First exploratory, then confirmatory (total $N = 126$, see our dataset package for details). These models yielded satisfactory fit indices and confirmed the oblique relationships among relevant items (e.g., values of Pledge Class Liking items differed from values of Active Liking items). Next, responses for survey items were combined into mean scores to create aggregated values for each measure, thus reducing 42 variables to

[3] In an effort to preserve participant anonymity, our public data file has been modified to exclude the combined mother's birthday/favorite color data field.

a set of nine variables for analysis.

For our focal analyses, we employed multi-level, or mixed-effects regression models to test for the predictive effects of Induction Harshness and Induction Fun on ratings of solidarity. Excepting the outcome variables, these models were identical with regards to predictors and nested structure. The seven multi-level models were structured as follows:

2.3.1. Level 1

- Individual ratings of the outcome variable. Each model included one of the seven measures of solidarity as an outcome variable.

2.3.1.1. Fixed effects.

- Individual ratings of Induction Harshness (grand mean-centered).
- Individual ratings of Induction Fun (grand mean-centered).
- Time Point of measurement (1, 2, 3, 4, 5).
- Interaction between centered Induction Harshness & Induction Fun.

2.3.2. Level 2

- Random intercept for Individual. This allows for the initial level of an outcome variable to vary and be unique to each participant.
- Random slope for Time Point. This allows for any effect of Time Point to differ across the sample.

2.3.3. Level 3

- Random intercept for Pledge Class. This allows for the initial level of the outcome variable to vary and be unique to each pledge class.

3. Results

Table 2 presents descriptive information, bivariate correlations, and internal consistency estimates for the nine measures used in this study. All of our aggregate measures showed sufficient internal consistency ($0.73 \leq \alpha \leq 0.95$). No demographic information was collected from participants, though we present the response/attrition rates for each of the six pledge classes across all time points in Table 1. We used mixed-effects models to appropriately account for the nested, or systematically shared, variance explained by our Level 2 and Level 3 variables, although we were not specifically testing for the multi-level effects of these models. Instead, our focus was on the individual-level effects of induction experiences (harshness and fun) as an explanation for reports of group solidarity. Tables 3 and 4 present full results for the mixed-effects models, and we discuss the results relevant to the solidarity macro theory below.

3.1. Pledge class-focused measures of solidarity

Table 3 presents results for the predictive effects of Induction Harshness and Induction Fun on pledge class-focused measures of solidarity. Broadly, Induction Harshness only explained a significant level of variance in one of the four outcome variables, Pledge Class Other-in-Self, at a small positive magnitude ($B = 0.11$). Two of the other pledge class-focused solidarity variables yielded similarly sized—though not conventionally significant—positive relationships with Induction Harshness. In contrast, Induction Fun was positively and significantly associated with measures of pledge class-focused solidarity, and all effects observed were larger, though small to moderate in size (*B*s from 0.23 to 0.32).

3.2. Active and chapter-focused measures of solidarity

Table 4 presents results for the predictive effects of Induction Harshness and Induction Fun on active and chapter-focused measures of solidarity. Induction Harshness explained no significant variation in any of the three outcome variables. However, as with pledge class-focused solidarity, Induction Fun was positively and significantly associated with active and chapter-focused solidarity, with observed effects a bit larger in size than those seen with pledge class-focused solidarity (*B*s from 0.27 to 0.41).

3.3. Interaction of induction harshness and induction fun

We also explored how the predictive effects of Induction Harshness may be conditioned by the Induction Fun experienced by pledges. As Table 3 reveals, the interaction between Induction Harshness was statistically significant for all measures of pledge class-focused solidarity. As an example of this effect, consider Pledge Class Communal Strength (Fig. 1). For pledges who reported low levels of Induction Fun (−1 *SD*), Induction Harshness had a positive slope with respect to Pledge Class Communal Strength ($B = 0.19$, $SE = 0.07$, $p = .01$). For pledges who reported average (0 *SD*) or high levels of Induction Fun (+1 *SD*), Induction Harshness had (respectively) a slightly positive non-significant slope ($B = 0.08$, $SE = 0.06$, $p = .17$ and a slightly negative non-significant slope ($B = -0.03$, $SE = 0.06$, $p = .60$). This same pattern of simple slope values held across all pledge class-focused solidarity variables.

Comparatively, none of the interaction terms between Induction Harshness and Induction Fun were statistically significant for the active and chapter-focused solidarity variables (Table 4).

Table 2
Descriptive information, bivariate correlations, and internal consistency estimates for measured variables.

Measured variable	*M*	*SD*	Range	Bivariate correlations[a]								
				1	2	3	4	5	6	7	8	9
1. Induction Harshness	4.21	1.39	1.00 – 7.00	0.90								
2. Induction Fun	5.08	1.20	1.00 – 7.00	−0.38*	0.90							
3. Pledge Class Liking	5.49	1.09	1.75 – 7.00	0.19*	0.26*	0.83						
4. Pledge Class Coordination	4.96	1.02	1.25 – 7.00	0.06	0.25*	0.64*	0.73					
5. Pledge Class Communal Strength	7.31	1.63	2.00 – 10.00	0.09	0.34*	0.68*	0.54*	0.95				
6. Pledge Class Other-in-Self	5.24	1.56	1.00 – 7.00	0.22*	0.21*	0.75*	0.51*	0.57*	–			
7. Active Liking	5.03	1.09	1.25 – 7.00	−0.03	0.45*	0.65*	0.48*	0.53*	0.49*	0.81		
8. Chapter Communal Strength	6.72	1.74	1.00 – 10.00	−0.04	0.46*	0.56*	0.42*	0.78*	0.42*	0.61*	0.95	
9. Chapter Other-in-Self	4.50	1.57	1.00 – 7.00	0.05	0.32*	0.60*	0.39*	0.49*	0.69*	0.68*	0.56*	–

Total $N = 126$.

[a] Cronbach's Alpha estimates of internal consistency are displayed in the diagonal for all multi-item measures.

* $p < .05$.

Table 3
Results of multi-level regression models for pledge class-focused solidarity variables.

	Outcome variables							
	Pledge Class Liking		Pledge Class Coordination		Pledge Class Communal Strength		Pledge Class Other-in-Self	
Predictors	Beta	95% CI	Beta	95% CI	Beta	95% CI	Beta	95% CI
(Intercept)	−0.07*	−0.19 – 0.04	−0.00*	−0.18 – 0.18	−0.02*	−0.15 – 0.12	−0.02*	−0.15 – 0.11
Time Point	0.50*	0.42 – 0.57	0.34*	0.23 – 0.45	0.26*	0.17 – 0.35	0.48*	0.40 – 0.57
Induction Harshness	0.08	−0.00 – 0.16	−0.01	−0.13 – 0.10	0.07	−0.03 – 0.16	0.11*	0.02 – 0.20
Induction Fun	0.23*	0.16 – 0.31	0.27*	0.17 – 0.37	0.32*	0.23 – 0.40	0.23*	0.15 – 0.32
Interaction Term[a]	−0.13*	−0.18 – −0.08	−0.08*	−0.15 – −0.01	−0.10*	−0.15 – −0.04	−0.07*	−0.12 – −0.01
Statistic	Random effects							
σ^2	0.27		0.45		0.74		0.66	
$\tau_{00\ Individual}$	1.03		0.76		2.09		2.06	
$\tau_{00\ Pledge\ Class}$	0.00		0.03		0.00		0.01	
$\tau_{11\ Individual\ Time\ Point}$	0.03		0.06		0.09		0.07	
$\rho_{01\ Individual}$	−0.92		−0.82		−0.65		−0.89	
ICC			0.48				0.58	
$N_{\ Pledge\ Class}$	6		6		6		6	
$N_{\ Individual}$	126		125		126		126	
Observations	399		386		400		396	
Marginal / Conditional R^2	0.614 / N/A		0.19 / 0.58		0.39 / N/A		0.34 / 0.72	

* $p < .05$.
[a] Interaction of centered Induction Harshness and centered Induction Fun.

Table 4
Results of multi-level regression models for active and chapter-focused solidarity variables.

	Outcome variables					
	Active Liking		Chapter Communal Strength		Chapter Other-in-Self	
Predictors	Beta	95% CI	Beta	95% CI	Beta	95% CI
(Intercept)	−0.03*	−0.16 – 0.11	0.00*	−0.16 – 0.16	0.04*	−0.12 – 0.19
Time Point	0.30*	0.21 – 0.38	0.18*	0.08 – 0.27	0.46*	0.37 – 0.55
Induction Harshness	−0.01	−0.10 – 0.09	0.00	−0.09 – 0.10	−0.01	−0.10 – 0.08
Induction Fun	0.39*	0.30 – 0.48	0.41*	0.33 – 0.50	0.27*	0.18 – 0.35
Interaction Term[a]	−0.04	−0.10 – 0.01	−0.05	−0.10 – 0.01	0.02	−0.03 – 0.08
Statistic	Random effects					
σ^2	0.39		0.82		0.62	
$\tau_{00\ Individual}$	0.64		1.59		2.13	
$\tau_{00\ Pledge\ Class}$	0.00		0.03		0.03	
$\tau_{11\ Individual\ Time\ Point}$	0.02		0.10		0.09	
$\rho_{01\ Individual}$	−0.60		−0.46		−0.79	
ICC	0.54		0.66		0.65	
$N_{\ Pledge\ Class}$	6		6		6	
$N_{\ Individual}$	126		126		126	
Observations	401		401		395	
Marginal / Conditional R^2	0.24 / 0.65		0.19 / 0.73		0.28 / 0.75	

* $p < .05$.
[a] Interaction of centered Induction Harshness and centered Induction Fun.

4. General discussion

There is no doubt that shared hardships can occasion human bonding. But not all hardships do so, and it is still questionable whether hazing typically manifests in a way that is highly conducive to such bonding. Thus far, it has proven surprisingly difficult to demonstrate—with measurement—that this is the case for real-world hazing groups. And while brief simulations of hazing in the laboratory have sometimes shown positive effects on solidarity, such results have been generated with ordeals that represent pale imitations of the intensity, duration, and context of many actual hazing processes.

Our study examined participants undergoing a real-world hazing induction process and assessed seven different measures of group solidarity. Our results are consistent with the possibility that hazing experiences are not substantially related to feelings of solidarity. For pledges to Beta, Induction Harshness evidenced only a small significant association with one measure of solidarity, Pledge Class Other-in-Self. Even if we ignore conventional significance, we are left with a harvest of small to trivial associations between Induction Harshness and measures of solidarity, associations substantially exceeded by those with Induction Fun (Tables 3 and 4). We also tested for interactions between Induction Harshness and Induction Fun. We found that a subset of pledges who reported particularly low levels of Induction Fun had significant but small associations between Induction Harshness and measures of pledge class-focused solidarity (Table 3). A plausible interpretation of this finding is that a subset of pledges who were having the *least* enjoyable induction experience sought and received more support from their pledge class, and in the process felt a bit more solidarity therewith. We do not think that these small and contingent effects are a great confirmation of the solidarity macro theory as it is typically expressed,

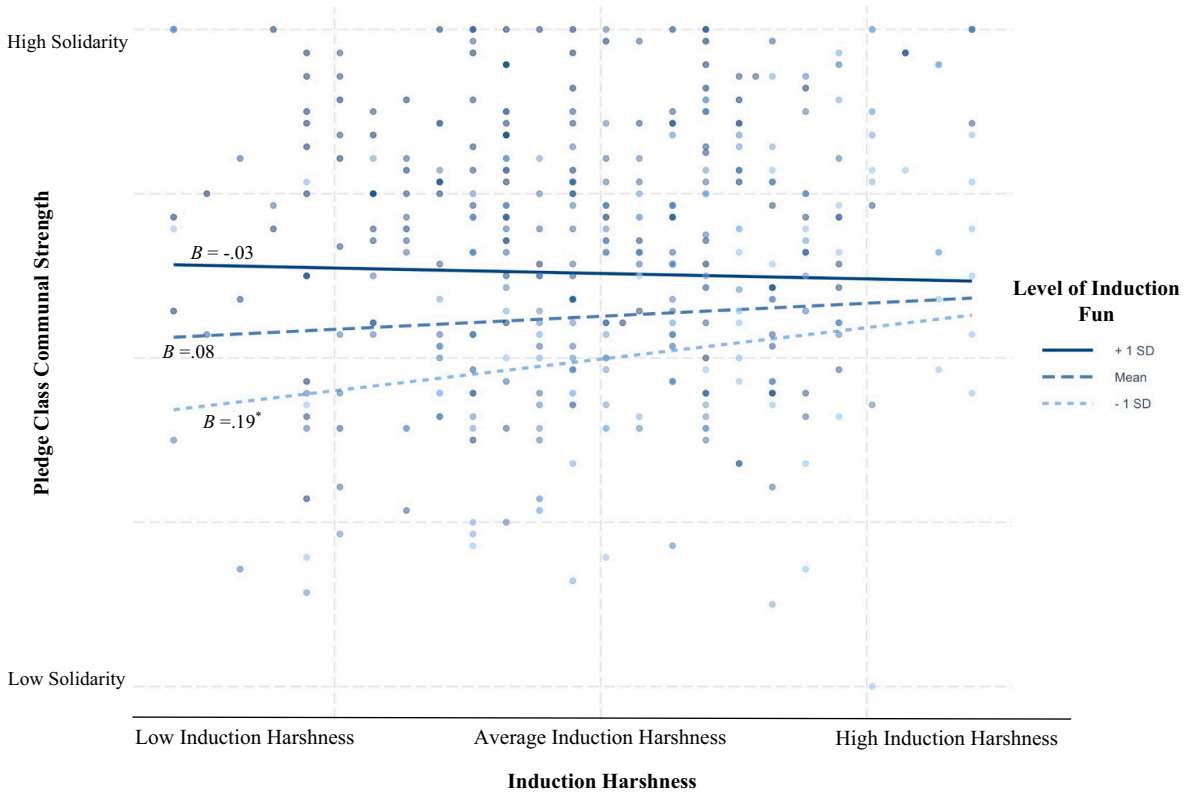

Fig. 1. Interaction of centered Induction Harshness and centered Induction Fun.
* $p < .05$

especially when compared to the effect sizes for Induction Fun by itself. Thus, our study is an important addition to the list of negative or equivocal findings for the solidarity macro theory. This is not just because of the overall pattern of our findings, but because of the conditions under which they were obtained: In situ, over time, from a real-world hazing group. To our knowledge, the only other researchers who have successfully done so are Lodewijkx and Syroit (1997), and they too generated findings inconsistent with the solidarity macro theory. Our study builds on Lodewijkx and Syroit (1997) by using a much broader set of solidarity measures, covering multiple incoming newcomer groups, and systematically accounting for variance at multiple levels. We also note that the internal consistency of our measures—when combined with their factor-analytic differentiation—suggest that participants were paying attention to our survey questions. This is additionally reflected in the observed relationships between Induction Fun and measures of solidarity, which are unlikely to result from random responding.

That said, there are a number of limitations to our results. First, given our sample size, it is possible that some real and non-trivial associations failed to manifest by chance alone. Further, it is possible that some aspect of our measurement strategy was insufficient. For example, identity fusion theorists suggest that fusing with one's group may require personal introspection on shared dysphoric experiences. Perhaps the participants in our study did not have enough time to reflect over their induction period (but see Mann et al., 2016, Study 1). Or perhaps our seven measures of solidarity are still missing a key operationalization that would reveal a substantial positive association. With respect to our induction measures, perhaps participants interpreted harsh hazing as "fun" (à la Keating et al., 2005). Some hazing ordeals may indeed have moments of levity, and (especially in the case of mild, brief hazing) may not be processed as primarily dysphoric (e.g., Thomas et al., 2021). Regardless, this kind of explanation becomes increasingly less likely as hazing ordeals are repeated over time and increase in severity. We find it implausible, for example, that most pledges would describe typical fraternity hazing ordeals as "pleasant", especially when made to continually endure them (e.g., Cimino, 2016).

Because we studied only one fraternity chapter, there may be some unknown property of this group that disrupted associations that would be present in other chapters. More broadly, we note that most of the relevant quantitative data on hazing has been derived from Western student populations. While this makes much of our literature review highly relevant to our studied sample, it impacts the general applicability of our results to hazing as a human phenomenon (Henrich, Heine, & Norenzayan, 2010). Such applicability depends on whether similarities in hazing across cultures at least partly reflect similar causal processes. This has long been regarded as plausible in the social sciences, and there is systematic evidence for this claim, but it remains a question under active exploration (e.g., Cimino, 2016; Cimino et al., 2019). Finally, as with all non-experimental studies, we cannot draw firm causal inferences from observed relationships.

All of these limitations are important to bear in mind, especially given that the science of hazing is still nascent and there are many unknowns. Indeed, our study should not be interpreted as settling the question of whether hazing tends to create solidarity. But we also want to warn against making the solidarity macro theory effectively non-falsifiable. Broadly speaking, we interpret the solidarity macro theory as predicting a substantial and obvious main effect, not a hard-to-detect, caveat-burdened association that reveals itself only to highly specific and precisely timed measurements. We think the solidarity macro theory has the potential to be particularly resistant to disconfirmation because the notion of hazing creating solidarity accords strongly with human intuitions about shared hardships, not to mention a rich body of supportive hazing anecdotes. But while anecdotes and intuitions can make for good starting points in the scientific process, they can also fill in gaps where more rigorous evidence would otherwise be expected. And when it comes to the science of hazing and group solidarity, such evidence is in

short supply. Thus, we suggest that it is time to reset our understanding of hazing's relationship with group solidarity. This does not mean rejecting the solidarity macro theory, this means treating it with a level of reserve that matches the rigor and consistency of its direct empirical support.

Moving forward, we think the relevant literature on the effects of hazing would benefit from expanding its current focus on complicated cognitive processes (e.g., cognitive dissonance, identity fusion). While these processes are worthy of consideration, they are being used to intricately explain an effect of hazing that is itself in question (i.e., solidarity). And in continually emphasizing such processes, we may end up underplaying the kind of mundane, unresolved ambivalence that hazees can experience (Allan, Kerschner, & Payne, 2019; Allan & Madden, 2008; Hoover & Pollard, 2000). To put this another way, while it is possible that many of the existing retrospective surveys on hazing have inconsistent findings because of validity issues, it is also possible that they are (roughly) reflecting something real: Hazing has mixed, unreliable effects on solidarity.

If hazing does have mixed effects on solidarity, why might this be so? One possibility is that lengthy, intense hazing processes are in many instances reflections of evolved motivations that were not designed to generate solidarity in newcomers. These motivations may instead be oriented toward cowing and controlling potentially exploitative newcomers and (where possible) encouraging the less-committed among them to disassociate (see Cimino et al., 2019). As a shared hardship, such a process might also create solidarity among newcomers, but it might do so in a way that is inefficient or inconsistent because it is not well-designed for that purpose. To clarify, imagine for a moment that there are possible ways of conducting hazing inductions that are highly effective at creating solidarity. The question is whether these hypothetical, solidarity-oriented hazing inductions look substantially similar to real-world hazing inductions. For example, Mann et al. (2016) argue that the common tactic of humiliating hazees may be particularly detrimental to feelings of solidarity. Similarly, requiring hazees to individually perform servile labor may be advantageous for hazers and their organization (Cimino, 2013b; Cimino et al., 2019), but of little benefit to hazee cohesion. Understanding the differing impacts of various kinds of hazing ordeals, and what proximate purposes (if any) they plausibly serve may be necessary in order to answer key theoretical and empirical questions. We hope that future studies will allow us to disentangle these issues and continue to unpack the nature of hazing.

Declaration of Competing Interest

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

References

Allan, E. J., Kerschner, D., & Payne, J. M. (2019). College student hazing experiences, attitudes, and perceptions: Implications for prevention. *Journal of Student Affairs Research and Practice, 56*(1), 32–43.

Allan, E. J., & Madden, M. M. (2008). Hazing in view: College students at risk. https://stophazing.org/research/national-hazing-study-hazing-in-view/.

Allan, E. J., & Madden, M. M. (2012). The nature and extent of college student hazing. *International Journal of Adolescent Medicine and Health, 24*(1), 83–90.

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596.

Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Applied Psychology, 45*(1), 55–58.

Barber, J. P. (2012). "Ever after strictly and rigidly obeyed-with some exceptions": Administrative responses to hazing in the 1870s. *Oracle, 7*(1), 1–21.

Cimino, A. (2007). Is hazing a byproduct? Testing the motivational systems that underlie the abuse of new group members. In *19th Human Behavior and Evolution Society Conference, Williamsberg, VA*.

Cimino, A. (2011). The evolution of hazing: Motivational mechanisms and the abuse of newcomers. *Journal of Cognition and Culture, 11*, 241–267.

Cimino, A. (2013a). Hazing as a manifestation of evolved psychology [dissertation, UC Santa Barbara]. https://www.proquest.com/docview/1448529589.

Cimino, A. (2013b). Predictors of hazing motivation in a representative sample of the United States. *Evolution and Human Behavior, 34*(6), 446–452.

Cimino, A. (2016). Fraternity hazing and the process of planned failure. *Journal of American Studies, 52*(1), 1–23.

Cimino, A. (2017). Defining hazing: Why popular definitions are misleading and counterproductive. *Journal of Higher Education Management, 32*(1), 135–148.

Cimino, A. (2020). Embrace & reform: Ending hazing's prohibition era. http://www.aldocimino.com/cimino_2020.pdf.

Cimino, A., Toyokawa, W., Komatsu, M., Thomson, R., & Gaulin, S. J. C. (2019). Cross-cultural and cross-organizational evidence for an evolved hazing motivation. *Evolutionary Psychology, 17*(4), 1–10.

Durkheim, É. (1912). *Les formes élémentaires de la vie religieuse: Le système totémique en Australie*. K. E. Fields, Trans.. *1995* Free Press (F. Alcan)

Ember, C. R., & Ember, M. (2010). Explaining male initiation ceremonies: New cross-cultural tests and a catalytic model. *Journal of Cross-Cultural Psychology, 41*(4), 605–616.

Enge, R. S. (1993). Effects of severity of initiation and feedback type on potential member attraction to an organization. *Dissertation Abstracts International: Section B. The Sciences and Engineering, 53*(9), 4989.

Gill, S. D. (1996). Disenchantment: A religious abduction. In R. L. Grimes (Ed.), *Readings in ritual studies* (pp. 230–239). Prentice Hall.

Grimes, R. L. (2000). *Deeply into the bone: Re-inventing rites of passage*. University of California Press.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *The Behavioral and Brain Sciences, 33*(2–3), 61–83.

Herdt, G. H. (1998). *Rituals of manhood: Male initiation in Papua New Guinea*. Transaction.

Hoover, N. C., & Pollard, N. J. (2000). Initiation rites in American high schools: A national survey. *Final report.*, 1–41. https://eric.ed.gov/?id=ED445809.

Kamau, C. (2013). What does being initiated severely into a group do? The role of rewards. *International Journal of Psychology, 48*(3), 399–406. https://doi.org/10.1080/00207594.2012.663957

Keating, C. F., Pomerantz, J., Pommer, S. D., Ritt, S. J. H., Miller, L. M., & McCormick, J. (2005). Going to college and unpacking hazing: A functional approach to decrypting initiation practices among undergraduates. *Group Dynamics: Theory, Research, and Practice, 9*(2), 104–126.

Lodewijkx, H. F. M., & Syroit, J. E. M. M. (1997). Severity of initiation revisited: Does severity of initiation increase attractiveness in real groups? *European Journal of Social Psychology, 27*(3), 275–300.

Lodewijkx, H. F. M., & Syroit, J. E. M. M. (2001). Affiliation during naturalistic severe and mild initiations: Some further evidence against the severity-attraction hypothesis. *Current Research in Social Psychology, 6*(7), 90–107.

Mann, L., Feddes, A. R., Doosje, B., & Fischer, A. H. (2016). Withdraw or affiliate? The role of humiliation during initiation rituals. *Cognition and Emotion, 30*(1), 80–100. https://doi.org/10.1080/02699931.2015.1050358

Mawritz, M. B., Capitano, J., Greenbaum, R. L., Bonner, J. M., & Kim, J. (2020). Development and validation of the workplace hazing scale. *Human Relations, 75*(1), 139–176, 0018726720957058.

McCreary, G., & Schutts, J. (2019). Why hazing? Measuring the motivational mechanisms of newcomer induction in college fraternities. *Journal of Cognition and Culture, 19*(3–4).

McMinn, B. L. (1980). A content analysis of the esoteric ritual manuals of national college social fraternities for men. *Dissertation Abstracts International: Section A. Humanities and Social Sciences, 40*(7), 3815.

Mills, J., Clark, M. S., Ford, T. E., & Johnson, M. (2004). Measurement of communal strength. *Personal Relationships, 11*(2), 213–230.

Owen, S. S., Burke, T. W., & Vichesky, D. (2008). Hazing in student organizations: Prevalence, attitudes, and solutions. *Oracle: The Research Journal of the Association of Fraternity/Sorority Advisors, 3*(1), 40–58.

Parks, G., & Brown, T. L. (2005). "In the fell clutch of circumstance": Pledging and the black Greek experience. In T. L. Brown, G. Parks, & C. M. Phillips (Eds.), *African American fraternities and sororities: The legacy and the vision* (pp. 437–464). University Press of Kentucky.

Precourt, W. E. (1975). Initiation ceremonies and secret societies as educational institutions. In R. W. Brislin, S. Bochner, & W. J. Lonner (Eds.), *Cross-cultural perspectives on learning* (pp. 231–250). Sage Publications.

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

Rogers, S., Rogers, C., & Anderson, T. (2012). Examining the link between pledging, hazing, and organizational commitment among members of a black greek fraternity. *Oracle: The Research Journal of the Association of Fraternity/Sorority Advisors, 7*(1), 43–53.

Schachter, S. (1959). *The psychology of affiliation: Experimental studies of the sources of gregariousness* (p. 141). Stanford Univer. Press.

Schopler, J., & Bateson, N. (1962). A dependence interpretation of the effects of a severe initiation. *Journal of Personality, 30*(4), 633–649.

Swann, W. B., Jetten, J., Gómez, Á., 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

Thomas, B. J., Cimino, A., & Meglich, P. (2021). Workplace hazing: Toward an organizational science of a cryptic group practice. *Group & Organization Management, 46*(2), 286–326. https://doi.org/10.1177/1059601121992893

Tiger, L. (1984). *Men in groups* (2nd ed.). Marion Boyars.

Turner, V. W. (1967). *The forest of symbols: Aspects of Ndembu ritual*. Cornell University Press.

Van Gennep, A. (1909). *Les rites de passage (M. B. Vizedom & G. L. Caffee, Trans.). 1960.* University of Chicago Press (É. Nourry).
Van Raalte, J. L., Cornelius, A. E., Linder, D. E., & Brewer, B. W. (2007). The relationship between hazing and team cohesion. *Journal of Sport Behavior, 30*(4), 491–507.
Whitehouse, H. (1996). Rites of terror: Emotion, metaphor and memory in Melanesian initiation cults. *Journal of the Royal Anthropological Institute, 2*(4), 703–715.
Whitehouse, H., Jong, J., Buhrmester, M. D., Gómez, Á., Bastian, B., Kavanagh, C. M., … Gavrilets, S. (2017). The evolution of extreme cooperation via shared dysphoric experiences. *Scientific Reports, 7*, 44292. https://doi.org/10.1038/srep44292
Whitehouse, H., & Lanman, J. A. (2014). The ties that bind us: Ritual, fusion, and identification. *Current Anthropology, 55*(6), 674–695. https://doi.org/10.1086/678698
Whiting, J. W. M., Kluckhohn, R., & Anthony, A. (1958). The function of male initiation ceremonies at puberty. In E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.), *Readings in social psychology* (pp. 359–370). Henry Holt & Company.
Young, F. W. (1965). *Initiation ceremonies: A cross-cultural study of status dramatization.* Bobbs-Merrill.