

Misinterpretation mechanisms and intergroup conflict following the mainstream diffusion of internet memes: a meme archaeology and translational action study based on ACG subculture

Zihan Lin

Beijing Haidian Kaiwen Academy, Beijing, China

banzhuren@yishan.org

Abstract. This study examines ten Internet memes originating from the Anime, Comics, and Games (ACG) subculture—"Gugu Gaga", "Hakimi", "Manbo", "Nailong", "He Yiwei", "Fei Yangyang", "Mu Guizi", "Sao Furry", "Feizhai", and "Yuanpi"—as research samples. Drawing upon meme theory, diffusion of innovations theory, and cultural reproduction theory, it investigates the mechanisms through which these memes are misinterpreted after entering the mainstream and the intergroup conflicts that subsequently arise. A mixed-methods approach was employed. A total of 490 valid questionnaires were collected, including responses from 159 core ACG community members, 271 individuals with limited exposure to the subculture, and 60 participants with no prior engagement. In addition, five semi-structured in-depth interviews were conducted. The findings indicate that the dissemination of Internet memes follows a structural logic of "meme entropy increase". Context-stripping dissemination is characterized by spontaneity, near-zero dissemination costs, and high communicative returns, whereas efforts to preserve or restore the original context require substantial investment while carrying a considerable risk of backlash. This structural asymmetry constitutes the primary driving force behind intergroup conflict. Regarding patterns of misinterpretation, three major forms were identified—character simplification, semantic shift, and malicious stigmatization—which together form a differentiated spectrum of cultural harm. Based on these findings, this study proposes a lightweight contextual supplementation strategy, the "Meme Interpretation Card", to facilitate more accurate understanding. Survey results show that 74.7% of respondents expressed a willingness to read such explanatory materials, while more than 80% agreed that public education about meme origins and meanings could effectively reduce disputes. These findings suggest that reconciliation between online communities enjoys broad public support and represents a practical and feasible direction for mitigating intergroup conflict.

Keywords: internet memes, ACG subculture, intergroup conflict, meme entropy increase, contextual restoration

1. Introduction: research background

Internet memes have become one of the core symbolic units of participatory online culture. Emerging and circulating within specific online communities, they not only embody the emotional identity of subcultural groups but are also extensively replicated, remixed, and frequently misinterpreted as they spread beyond their original communities into the mainstream. In recent years, the cross-community diffusion of numerous memes originating from the Anime, Comics, and Games (ACG) subculture has generated substantial intergroup conflict and attracted widespread public attention. This raises an important question worthy of systematic investigation: Why do memes that are highly valued within their original subcultural communities often provoke strong negative reactions after entering the mainstream? What mechanisms underlie such misinterpretations? More importantly, is there a viable pathway for mitigating these conflicts?

This study addresses three central research questions: Why are Internet memes misinterpreted after they diffuse beyond their original communities? What underlying mechanisms link these misinterpretations to intergroup conflict? And is there a feasible pathway toward reconciliation? To answer these questions, the study selects ten representative Internet memes that originated within the ACG subculture and were subsequently subject to varying degrees of misinterpretation or stigmatization: "Gugu Gaga", "Hakimi", "Manbo", "Nailong", "He Yiwei", "Fei Yangyang", "Mu Guizi", "Sao Furry", "Dead Fat Otaku", and "Yuanpi". It should be noted that all selected cases had already generated noticeable discomfort or conflict within the ACG community. This purposive sampling strategy enables an in-depth examination of the mechanisms underlying intergroup conflict. However, it also limits the generalizability of the findings, as they should not be interpreted to suggest that every meme entering the mainstream will inevitably produce similar forms of harm. Methodologically, this study adopts a mixed-methods approach, combining quantitative survey research with qualitative interviews. A total of 490 valid questionnaires were collected, covering three participant groups: 159 core members of the ACG community, 271 individuals with limited exposure to the subculture, and 60 participants with no prior engagement with the community. In addition, five semi-structured in-depth interviews were conducted with participants representing four groups: general Internet meme users, core ACG community members, individuals entirely outside the subculture, and older-generation participants with no involvement in the community. The interview data were analyzed using thematic coding, and the resulting themes serve as supplementary evidence supporting the quantitative findings.

2. From an "inside joke" to an overused mainstream meme

2.1. The life cycle of an internet meme

Before examining specific cases, this study proposes a conceptual framework for understanding the life cycle of Internet memes. Across the cases analyzed, memes exhibit a remarkably consistent evolutionary trajectory: they begin as shared symbols within a small community, carrying well-defined meanings and emotional significance; after crossing into the mainstream, they undergo contextual dilution and a shift in interpretive authority, eventually becoming overused memes—that is, expressions employed so broadly that they become detached from their original cultural context. It should be noted that the term "overused" refers to a pattern of dissemination rather than a value judgment about a meme's quality. The transformation is driven primarily by the dynamics of information diffusion rather than by the moral failings of any particular group. However, in cases involving deliberate stigmatization, those who intentionally initiate malicious reinterpretations should bear corresponding ethical responsibility.

2.2. Origins within a small community

When a meme first emerges, it typically possesses a clear community affiliation and well-defined cultural meaning. Its value lies not in profound semantic depth but in its shared ownership by the community that created it. Participants understand the meme intuitively because they share the same cultural background, making explicit explanation unnecessary. For example, within the "*MyGO!!!!*" fandom, "Gugu Gaga" originally functioned as an insider joke. Those who used it already knew who Takamatsu Tomori was, understood her association with penguins, and recognized that the distinctive voice clip originated from a fan-made remix rather than the original anime. Similarly, among fans of *Uma Musume: Pretty Derby*, "Manbo" refers to an adorable humming performed by the character Matikanetannhauser, instantly evoking a shared memory of a specific scene. Likewise, "He Yiwei" originated among enthusiasts of pseudo-Chinese writing as a linguistic game: the Japanese phrase "Nani no imi (何の意味)" is presented only in kanji, omitting the kana, thereby producing a humorous expression that appears Chinese while remaining distinctly Japanese. Within these relatively closed communities, the meanings of such memes are stable and unambiguous. They function much like badges of membership: those who use them immediately recognize one another, and misunderstandings rarely occur because the original cultural context remains intact.

2.3. The cost of going mainstream

The transition into mainstream circulation marks a turning point in a meme's life cycle. A meme may spread widely because of its absurdity, catchiness, or striking contrast. Once reposted onto larger platforms, recommendation algorithms amplify its visibility, exposing it to countless users who have little or no knowledge of its original source. These audiences encounter only isolated fragments—a short clip, a voice recording, or a facial expression—rather than the broader cultural context from which the meme emerged. The case of "Gugu Gaga" illustrates this process. The meme originated from a fan-edited video featuring Takamatsu Tomori from the anime "*It's MyGO!!!!*". Owing to the humorous contrast created in the edit, the clip rapidly gained popularity, accumulating more than 17.37 million views on Bilibili. Yet the overwhelming majority of those who subsequently circulated the meme were unaware of Tomori's original characterization or the fan-made origins of the audio. As a result, the character was reduced to the simplistic stereotype of a "penguin girl", effectively becoming a reaction image stripped of her original personality and narrative depth.

Survey findings further reveal the structural characteristics of meme dissemination. Among all respondents, 44.8% reported that they do not actively seek information about a meme's original source, while 60.4% agreed that "a meme only needs to be useful or funny". Regarding exposure channels, 95.1% indicated that they primarily encounter new memes passively through platforms such as Douyin and Bilibili, while more than 60% reported learning about memes through everyday interactions with friends and classmates. Passive algorithmic recommendation, combined with interpersonal social transmission, thus constitutes the principal pathway through which Internet memes spread and, simultaneously, the primary breeding ground for misinterpretation [1]. This observation is consistent with previous scholarship. Dawkins [2] argues that cultural units possessing greater replicability are more likely to spread successfully. Sweller's Cognitive Load Theory [3] suggests that individuals naturally prefer information requiring less cognitive effort to process. Likewise, Rogers' Diffusion of Innovations Theory [4] emphasizes simplicity as one of the critical attributes facilitating the rapid diffusion of new ideas. Consequently, memes requiring less explanation and less background knowledge tend to spread more efficiently—not because users are inherently malicious, but because such memes align with fundamental principles of human cognitive economy and communicative efficiency.

2.4. Inevitable divergence

Once a meme enters mainstream circulation, its future generally follows one of several trajectories. Some memes retreat into their original communities after periods of controversy, as community members actively reclaim interpretive authority and the meme's circulation gradually contracts. Many others simply fade from public attention, disappearing as new memes replace them in the rapidly evolving online environment. A small number of memes undergo reappropriation. For instance, "Mu Guizi", originally coined as a derogatory label targeting "MyGO!!!!" fans, has been reclaimed by some members of the targeted community through self-deprecating humor, thereby weakening its original offensive intent. Finally, a handful of memes become fully integrated into everyday language. Expressions such as "bailan" ("to give up and coast") and "pofang" ("to have one's emotional defenses broken"), now widely used in ordinary conversation, were themselves once context-specific Internet memes rooted in particular online communities. Regardless of which trajectory a meme ultimately follows, one fundamental pattern remains consistent: mainstream diffusion inevitably dilutes its original meaning. This process should not be understood as a decline in quality but rather as a transformation into overuse—the meme becomes so widely and indiscriminately employed that it escapes the cultural boundaries within which it originally derived its significance, overflowing into contexts far removed from its source.

3. Meme entropy increase—the structural logic of online communication

The preceding discussion explains how Internet memes become misinterpreted after entering the mainstream. A more fundamental question, however, remains: Why is this process so difficult to prevent? Why is preserving the original context of a meme so challenging?

3.1. Translating the law of entropy into the internet context

The Second Law of Thermodynamics states that, within an isolated system, spontaneous processes move from order toward disorder, resulting in an increase in entropy. Conversely, restoring order requires a continuous input of external energy; once this effort ceases, the system naturally returns to disorder [5]. Although the online public sphere is not an isolated physical system, its dynamics of contextual preservation and contextual loss exhibit a strikingly similar pattern. Online anonymity weakens social constraints, low participation barriers reduce the cost of engagement, and recommendation algorithms systematically reward content that is emotionally stimulating and easily replicated. As a result, simplified memes requiring minimal cognitive effort spread rapidly, whereas clarifying and preserving their original meanings demands sustained interpretive labor. Previous studies have documented this structural asymmetry from multiple perspectives. Dawkins' meme theory [2] explains how cultural units with greater replicability achieve wider dissemination. Milner [6], in his study of participatory meme culture, demonstrates how memes continually evolve through collective remixing and reinterpretation. Meanwhile, Massanari's research on Reddit [7] illustrates how platform architecture and algorithmic systems amplify low-cost, emotionally driven forms of participation.

Building upon these studies, the present research proposes the heuristic framework of "meme entropy increase". It should be emphasized that this framework is not a literal application of thermodynamic principles. Rather, entropy serves as an analytical metaphor for describing the structural asymmetry between meme diffusion and contextual preservation. Its central proposition is that, in the dissemination of controversial memes, the costs of contextual distortion and simplification are substantially lower than those required for contextual restoration. Consequently, in online environments where guidance and intervention are generally limited, contextual loss typically outpaces contextual preservation. Compared with previous research, this

framework contributes in three respects. First, it explicitly links context stripping to entropy-increasing dissemination, proposing a life-cycle model that traces memes from insider memes, to overused mainstream memes, and eventually to ordinary everyday expressions. Second, drawing on survey and interview data, it provides empirical evidence demonstrating the high costs and substantial backlash risks associated with entropy-reducing behaviors. Third, the framework is applied to an in-depth analysis of ten controversial Internet memes, thereby evaluating its explanatory power at the micro level.

3.2. Meme usage as a spontaneous entropy-increasing process

In this study, entropy-increasing behavior refers to communicative practices such as reposting memes without context, imitating popular meme usage, and circulating memes while stripping away information about their original source. These behaviors share three defining characteristics: low cognitive cost, high emotional reward, and no requirement for additional effort. First, using a meme requires very little cognitive investment. Users need not investigate its origin, understand the original work, or worry about being challenged for inaccurate usage. Emotionally, they are under no obligation to empathize with others, exercise restraint, or assume responsibility for repairing misunderstandings. Moreover, online anonymity largely shields users from the consequences of inappropriate meme use. These characteristics closely resemble the spontaneous, energy-free nature of entropy increase in physical systems. Second, meme usage generates immediate positive reinforcement. As one interview participant remarked: "Posting a 'Gugu Gaga' meme is effortless. It is something you can do casually, and it helps you relax". This observation helps explain why meme dissemination is fundamentally self-propelling. Each time a meme is used, the user receives a small emotional reward. This continuous cycle of positive reinforcement accelerates diffusion without requiring any external driving force. By contrast, entropy-reducing behaviors—such as citing a meme's original source, explaining its intended meaning, introducing the background of the original work, or respecting contextual boundaries—display the opposite characteristics. They require substantial investments of time and emotional effort, offer little immediate reward, and frequently expose those engaging in them to ridicule or hostility. During the interviews, most core community members expressed reluctance to actively explain meme origins because doing so was both exhausting and likely to attract criticism. These findings reveal a striking structural asymmetry. Damaging a meme's original context is a spontaneous activity that provides immediate gratification at virtually no cost. Restoring that context, however, constitutes an entropy-reducing process that demands considerable effort while carrying a substantial risk of backlash.

3.3. Catalysts of intergroup conflict

It should be emphasized that differences in understanding alone do not inevitably escalate into large-scale online conflict. While indiscriminate meme usage represents entropy-increasing behavior, the actors who frequently transform misunderstandings into full-scale confrontations are a particular group of Internet users commonly referred to in Chinese online culture as "shenren" ("chaos instigators") or "leziren" ("spectators who enjoy stirring up drama"). These individuals derive enjoyment from provoking conflict and creating disorder. They deliberately distort the meanings of memes to antagonize different communities, finding entertainment in the resulting disputes. One interviewee from the ACG community described their behavior as follows: "They sit in the middle and provoke both sides. It is not that they do not understand the original meaning—in many cases, they understand it better than anyone else. They deliberately use distorted meanings to provoke one group, then enjoy watching the other group become angry". The dissemination of the "Sao Furry" meme clearly illustrates this catalytic role. Through the combined actions of disruptive insiders and

conflict-seeking provocateurs, a subculture that originally carried positive meanings gradually became subject to systematic public stigmatization.

3.4. Implications of the meme entropy increase framework

In "*What Is Life?*", Schrödinger [8] famously argued that life feeds on negative entropy. The universe naturally tends toward increasing disorder, yet living systems maintain local order through the continuous expenditure of energy. The same insight offers a useful perspective on today's online environment. It is unrealistic to expect the Internet as a whole to evolve spontaneously into an orderly communicative space. What individuals can do, however, is continuously invest effort within the limited spaces they can influence, thereby sustaining small pockets of order, understanding, and mutual respect. Explaining the origin of a meme after it goes mainstream, resisting the impulse to join indiscriminate meme trends, or patiently providing context to someone genuinely willing to listen—each of these seemingly insignificant actions represents a form of entropy reduction. Collectively, such efforts constitute acts of goodwill that help preserve cultural understanding in an increasingly disordered online environment.

4. Meme archaeology and the reconstruction of misinterpretation

Drawing on a meme archaeology of ten controversial memes originating from the ACG subculture, together with survey and interview data, this chapter provides empirical support for the meme entropy increase framework. Based on the dominant patterns of misinterpretation, the ten memes can be classified into three categories.

4.1. Type I: character simplification

Representative examples of this category include "Gugu Gaga" and "Fei Yangyang". Their common characteristic is that originally complex, emotionally rich fictional characters are reduced to a single exaggerated or highly memorable trait. Through repeated circulation on short-video platforms, they become flattened into one-dimensional symbols, while the emotional depth valued by their original fan communities is almost entirely erased.

The "Gugu Gaga" meme originated from Takamatsu Tomori, a character in the anime "*It's MyGO!!!!*". In the original work, Tomori is portrayed as a thoughtful, introverted girl who struggles with verbal expression and quietly collects what she considers "ordinary things". Her characterization reflects the loneliness, sensitivity, and quiet kindness often associated with introverted personalities. In January 2025, the Bilibili content creator Tushan Baibai Miao uploaded a fan-edited video combining Tomori's footage with an absurd vocal clip repeatedly chanting "Gugu Gaga". The video has since accumulated more than 17.37 million views, rapidly transforming the phrase into a viral Internet meme. Survey data reveal the cognitive consequences of this simplification. When respondents were asked whether they understood the origin of "Gugu Gaga", 78.6% of core ACG community members reported complete familiarity with its background, whereas only 3.2% of participants entirely outside the community gave the same response.

The case of "Fei Yangyang" follows a similar trajectory. Originally portrayed in the *Pleasant Goat and Big Big Wolf* franchise as a brave, righteous, and devoted protector of his companions, the character was reduced to isolated scenes depicting his unrequited affection. These fragments became the basis for labeling him a "simp", turning the character into a widely used insult for mocking others online. The defining harm of this category lies in the loss of character integrity. Fictional characters cease to exist as multidimensional

individuals with coherent personalities and narratives, becoming instead simplistic symbols representing a single emotion or stereotype.

4.2. Type II: semantic shift

Representative examples of this category include "Nailong", "He Yiwei", "Hakimi", and "Manbo". Compared with character simplification, semantic shift represents a more profound transformation: the original meaning of a word, phrase, or image is not merely reduced but fundamentally replaced, so that it comes to signify something entirely different.

The evolution of "He Yiwei" provides a representative example. The expression derives from the Japanese phrase "何の意味" ("nani no imi", "what does it mean?"), rendered in pseudo-Chinese by retaining only the Chinese characters while omitting the kana. Originally, it conveyed a sincere expression of curiosity or confusion. Over time, however, the phrase gradually acquired connotations of provocation, sarcasm, and baiting. What began as a neutral question evolved into a versatile expression of ironic mockery. Owing to its brevity and high replicability, it spread rapidly across online communities. Following its widespread adoption by fans of "*BanG Dream! It's MyGO!!!!*" in 2025, the phrase entered broader ACG discourse, where its meaning became increasingly fragmented. Depending on context, it may now function as lighthearted teasing, a cold interrogation, or simply meaningless repetitive spam.

The case of "Nailong" illustrates an even more extreme form of semantic transformation. This meme emerged through the forced fusion of three entirely unrelated sources: the animated children's character from "Nailong", the exaggerated facial expressions of content creator Ji Hongxing, and the distorted laughter of game streamer Bajige. Through AI-assisted editing and voice synthesis, Internet users stitched these disparate elements together into a hybrid creation that bears little resemblance to any of its original sources. Detached from all previous contexts, "Nailong" has become a self-contained symbol of exaggerated, infectious laughter. Interestingly, its dissemination also involved an element of deliberate provocation toward the original creators. Users flooded the official "Nailong" social media accounts with meme images and comments, eventually forcing the account administrators to disable the comment section. This episode exemplifies a form of deconstructive carnival, in which playful reinterpretation evolves into collective disruption.

The diffusion of "Hakimi" and "Manbo" followed a particularly distinctive trajectory. "Hakimi" originates from the Honey Song, hummed by the character Tokai Teio in the anime *Uma Musume: Pretty Derby*. The original lyric refers simply to a honey-flavored beverage and has no connection whatsoever with cats. "Manbo", by contrast, originates from another character in the same franchise, Matikanetannhauser, who performs a cheerful humming sequence during her personal story. In November 2022, the Bilibili creator Jingqiao Shana uploaded a parody song entitled *Two Fools*, widely regarded as the beginning of the so-called "Great Hakimi Era". Thereafter, countless creators overlaid the "Hakimi" soundtrack onto videos featuring adorable cats, leading millions of viewers to develop the mistaken association that "Hakimi" simply means "cat". The development of AI voice synthesis further accelerated this transformation. By extracting the original voices of Tokai Teio and Matikanetannhauser, creators produced voice models capable of singing nonsensical syllables such as "Hakimi", "Manbo", and "Nanbei Lüdou", generating highly addictive parody songs. In September 2024, the viral popularity of videos featuring the so-called round-headed elderly cat further reinforced the association between "Hakimi" and the image of an uncontrollable stray cat. Today, on Douyin alone, content associated with "Hakimi" has accumulated more than 40.42 billion views. When contemporary audiences hear the melody, it no longer signifies honey, cats, or even any stable semantic content. Instead, it functions primarily as a source of immediate auditory pleasure—a form of absurd, post-semantic cyber carnival in which emotional stimulation has completely displaced original meaning.

4.3. Type III: malicious stigmatization

Representative examples of this category include "Mu Guizi", "Sao Furry" (later transformed into "Sao Furry" as a homophonic variant), "Feizhai" (originally "Dead Fat Otaku"), and "Yuanpi". Unlike the previous two categories, these expressions did not originate as ordinary Internet memes. Rather, they began as intentionally derogatory labels and only became widely recognized after entering mainstream online discourse, where they proliferated on a much larger scale.

The term "Feizhai" ("fat otaku") originally emerged as a form of self-deprecating humor within the ACG community. A contraction of "overweight" and "otaku", it was commonly used on early platforms such as Tieba and Bilibili to jokingly describe a lifestyle characterized by staying indoors, playing video games, and avoiding exercise. In this context, self-mockery functioned as a strategy for coping with external stereotypes and prejudice. With the rise of short-video platforms, however, "Feizhai" was appropriated by users outside the community. Its pragmatic function shifted from self-directed humor to external ridicule. Individuals perceived as overweight, lacking self-discipline, socially awkward, or simply interested in ACG culture increasingly became targets of the label. The term was frequently combined with insults such as "disgusting" or "social underclass", transforming it into an instrument of personal humiliation rather than playful self-identification.

The term "Yuanpi" represents a systematic form of stigmatization directed at gaming communities in the Chinese Internet. It first emerged during conflicts surrounding the player community of the game *Genshin Impact*. Initially, opponents employed the label to criticize what they viewed as extreme or toxic fans. Very quickly, however, the criticism shifted from condemning particular behaviors to stigmatizing an entire identity. Simply playing *Genshin Impact* became sufficient for an individual to be labeled a "Yuanpi". Subsequently, the "X-pi" naming pattern evolved into a productive template applied indiscriminately to players of numerous other popular games.

The trajectory of "Mu Guizi" is more complex. The expression is a homophonic substitution for "Mu Guizi" ("female devils"), a derogatory term that originated within the Douyu "abstract" online community. In response to this slur, fans of "*MyGO!!!!*" gradually shifted from anger and resistance toward self-reappropriation. Some core community members intentionally adopted "Mu Guizi" as a self-designation for themselves and the works they admired. This process resembles what has been described as stigma reclamation [9], whereby the widespread voluntary adoption of a derogatory label within the targeted community diminishes its offensive power.

Among all cases examined, "Sao Furry" most clearly illustrates the process of meme entropy increase. Furry culture originally developed as a positive subculture centered on original artwork and anthropomorphic animal characters. However, following the creation of a small number of vulgar works by disruptive insiders and their extensive circulation by conflict-seeking Internet users, anti-Furry groups deliberately coined the derogatory label "Sao Furry" in early 2026. This label systematically associated the entire community with vulgarity and subsequently generated the aggressive slogan "*Sweep Away the Furies*". One respondent from the furry community wrote in the open-ended section of the survey: "Memes that originally carried a playful tone quickly deteriorate into hostile language, making the online environment increasingly toxic". Taken together, these examples demonstrate that Type III memes closely follow the meme entropy increase model. Creating and spreading stigmatizing labels requires virtually no effort, whereas those targeted must expend substantial time, emotional energy, and resources defending themselves and correcting public misconceptions.

4.4. Differences among the three types of misinterpretation

The three forms of misinterpretation identified in this study constitute a continuum ranging from relatively mild to highly damaging. At one end lies character simplification, represented by "Gugu Gaga" and "Fei

Yangyang". Here, multidimensional fictional characters are reduced to one-dimensional symbols, resulting in a moderate degree of cultural harm. The second category, semantic shift, represented by "Nailong", "He Yiwei", "Hakimi", and "Manbo", involves the complete replacement of original meanings, causing the disappearance of the source context and producing more substantial cultural distortion. The final category, malicious stigmatization, represented by "Mu Guizi", "Sao Furry", "Feizhai", and "Yuanpi", transforms memes into instruments of hostility through the systematic stigmatization of entire communities. This form produces the most severe social consequences. Although these categories differ significantly in the degree of harm they generate, they share the same underlying mechanism of dissemination—meme entropy increase. Consequently, while they can be analyzed within a common theoretical framework, their social impacts and appropriate responses should be clearly distinguished.

4.5. A comparison with benign mainstream memes

Not every meme that enters the mainstream produces conflict. Several widely circulated Internet expressions illustrate that cross-community diffusion can occur without generating substantial misunderstanding or hostility. The expression "YYDS" ("the greatest of all time") spread from the esports community into general online discourse while largely preserving its original meaning, generating little intergroup conflict. Likewise, "Pofang" ("to have one's emotional defenses broken") evolved naturally from a gaming term into an everyday expression describing emotional vulnerability, with its semantic broadening widely accepted by Internet users. Similarly, "Da Call", originally derived from Japanese idol-support culture, entered mainstream Chinese usage without significant distortion of its core meaning. These cases suggest that mainstream diffusion alone is not sufficient to produce cultural harm. Rather, harm depends on the manner in which contextual information is stripped away. When contextual loss is accompanied by deliberate distortion or malicious reinterpretation, intergroup conflict is likely to emerge. By contrast, when semantic broadening occurs organically through ordinary language change, integration between communities becomes possible. Distinguishing benign semantic evolution from malicious contextual stripping is therefore essential for understanding, and ultimately mitigating, conflicts arising from the mainstream diffusion of Internet memes.

5. Toward reconciliation between online communities

If entropy increase is a spontaneous tendency, then entropy reduction is a conscious choice.

5.1. The meme interpretation card

This study proposes the Meme Interpretation Card, a minimalist translational tool designed to provide essential contextual information through a simple image accompanied by concise explanatory text. The idea is inspired by the work of the Bilibili educational creator Meme Encyclopedia and by collaborative knowledge platforms such as Moegirlpedia, whose content is produced from the perspective of dedicated community members. For a meme such as "Gugu Gaga", a single card requiring only a few seconds to read could allow users outside the ACG community to learn, perhaps for the first time, that Takamatsu Tomori is a thoughtful and emotionally complex anime character rather than merely the source of a humorous sound clip. Such minimal contextual supplementation would likely prevent many cases of unintentional misuse. The Meme Interpretation Card also requires very little effort from either side. Community members can simply keep the card on their mobile devices and, when confronted with misuse, respond by sharing the card rather than engaging in prolonged online arguments. Likewise, for well-intentioned users outside the community, reading the card represents a virtually cost-free opportunity to acquire the missing contextual knowledge.

Survey results provide empirical support for this proposal. Across all respondent groups, 74.7% indicated that they would be willing to read a Meme Interpretation Card explaining a meme's origin, intended meaning, and appropriate context of use. Furthermore, more than 80% agreed that learning about a meme's background could effectively reduce unnecessary disputes. Interview responses reveal similar attitudes. One participant outside the ACG community commented: "People should be more understanding. If someone uses a meme incorrectly, it is probably because they do not know enough about it. They are not part of the community, so a little tolerance is appropriate". Meanwhile, interviewees from within the community consistently emphasized that their objective was not to prohibit others from using memes, but rather to encourage respect for the original context from which those memes emerged. These findings reveal the central tension underlying many intergroup conflicts. Users outside the community seek a low-cost way to enjoy Internet memes, while community members hope to preserve respect for the original works and their cultural meanings. Their disagreement therefore stems not from fundamentally different moral values but from a practical question: Who should bear the cost of bridging the information gap? The Meme Interpretation Card is proposed precisely as a response to this structural dilemma.

5.2. A multifactorial explanation of intergroup conflict

Before discussing possible solutions, it is important to clarify that escalating conflicts between online communities cannot be attributed solely to the malicious behavior of "drama-seekers" or "chaos instigators". Rather, such conflicts emerge through the interaction of multiple structural factors. First, platform recommendation algorithms systematically prioritize content characterized by conflict, emotional intensity, and strong entertainment value, thereby amplifying social polarization. Second, traffic-driven commercial incentives encourage creators to strip away contextual information, manufacture controversy, and stimulate online disputes in pursuit of greater visibility and engagement. Third, conflict-seeking users deliberately exploit these algorithmic dynamics, transforming ordinary differences in understanding into confrontations between communities. Finally, information asymmetry between communities provides the essential conditions under which conflict develops. It is precisely at this point that innocent misunderstanding and deliberate provocation converge, allowing relatively minor differences to escalate into widespread hostility. Consequently, any realistic strategy for reconciliation cannot rely on eliminating conflict-seeking individuals altogether. Instead, it should focus on reducing the information gap between communities—the factor that is both the most manageable and the primary concern of the present study.

5.3. Translational action as incremental social repair

Viewed from this perspective, the Meme Interpretation Card represents more than a practical communication tool. It also serves as a metaphor for translation itself. Within an online communication system characterized by accelerating information flows, increasing social fragmentation, and continual meme entropy increase, carefully constructed cultural contexts are gradually dispersed into isolated symbols detached from their original meanings. Under such circumstances, the objective should not be to erect barriers between communities but to build bridges connecting them. The Meme Interpretation Card functions as a means of restoring the contextual "birth certificate" that a meme has lost during its dissemination. Although this intervention may appear modest, it has tangible social value. It supports subcultural communities that have become targets of stigmatization while simultaneously assisting well-intentioned users outside those communities who are constrained primarily by informational asymmetry rather than hostility. More fundamentally, it promotes a constructive model of communication founded on two simple principles: understand before using, and listen before judging.

6. Conclusion and implication

6.1. Main findings

Through an in-depth analysis of ten controversial Internet memes originating from the ACG subculture, this study demonstrates the applicability of the meme entropy increase framework to the dissemination of online subcultural content and yields three principal findings. First, the diffusion of Internet memes follows the internal logic of meme entropy increase. As memes enter mainstream circulation, the dilution of their original meanings is not a matter of deterioration in quality but rather a consequence of overextension through excessive dissemination. Behaviors that strip away contextual information are characterized by spontaneity, negligible dissemination costs, and immediate communicative rewards, whereas efforts to preserve or restore context require substantial investments of time and emotional labor while exposing participants to a significant risk of backlash. This structural asymmetry provides the key to understanding the emergence of intergroup conflict. Second, the three categories of misinterpretation identified in this study form a differentiated spectrum of cultural harm. Although character simplification, semantic shift, and malicious stigmatization all operate according to the same underlying mechanism of meme entropy increase, they differ substantially in both their mechanisms of harm and their severity. Consequently, strategies for addressing these forms of misinterpretation should also be differentiated rather than treated uniformly. Third, the proposed Meme Interpretation Card receives empirical support from both the survey and interview findings. More than 70% of respondents expressed a willingness to read contextual interpretation cards, while over 80% agreed that increasing public understanding of a meme's origin and intended meaning could reduce unnecessary disputes. These findings suggest that meaningful reconciliation between online communities enjoys broad public support and represents a practically achievable objective.

6.2. Academic contributions

This study integrates insights from meme theory, participatory meme culture research, and scholarship on the increasing disorder of online communication to propose both a meme life-cycle model and operational definitions of entropy-increasing and entropy-reducing behaviors. Together, these contributions provide a theoretically grounded and empirically testable analytical framework for interdisciplinary research on Internet memes. Furthermore, by incorporating the perspectives of participants within ACG communities, the study enriches existing research on conflicts between online communities with empirical evidence derived from an insider perspective, thereby complementing previous studies that have primarily emphasized external observations.

6.3. Practical implications

The Meme Interpretation Card offers a lightweight approach to contextual supplementation that may serve as a practical and low-cost tool for mitigating conflicts between online communities. Beyond its application to ACG culture, the approach has potential value in a range of settings, including community-based public education, family media literacy education, and platform content governance, where improving contextual understanding may contribute to more constructive online interactions.

6.4. Limitations

Several limitations should be acknowledged. First, the meme entropy increase framework proposed in this study is primarily intended to explain the dissemination of controversial memes that have already generated

intergroup conflict. Whether the framework is equally applicable to the diffusion of benign mainstream memes remains an open question that should be examined through comparative case studies in future research. Second, the empirical findings are based on 490 questionnaire responses and five semi-structured interviews. Consequently, the conclusions should be interpreted as the results of an exploratory investigation rather than as statistically generalizable inferences. Third, the practical effectiveness of the Meme Interpretation Card has yet to be evaluated through implementation in real-world communication environments. Its capacity to reduce misunderstanding and conflict therefore remains an empirical question requiring further investigation. Nevertheless, the significance of this study lies not in claiming to provide definitive answers, but in offering a theoretically informed and empirically grounded starting point for understanding the mechanisms underlying meme misinterpretation and for exploring constructive approaches to fostering dialogue across online communities.

References

- [1] Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.
- [2] Dawkins, R. (1976). *The selfish gene*. Oxford University Press.
- [3] Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- [4] Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- [5] Halliday, D., Resnick, R., & Walker, J. (2013). *Fundamentals of physics*. John Wiley & Sons.
- [6] Milner, R. M. (2018). *The world made meme: Public conversations and participatory media*. MIT Press.
- [7] Massanari, A. (2017). #Gamergate and the fapping: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media & Society*, 19(3), 329–346.
- [8] Schrödinger, E. (1944). *What is life? The physical aspect of the living cell*. Cambridge University Press.
- [9] Galinsky, A. D., Wang, C. S., Whitson, J. A., Anicich, E. M., Hugenberg, K., & Bodenhausen, G. V. (2013). The reappropriation of stigmatizing labels: The reciprocal relationship between power and self-labeling. *Psychological Science*, 24(10), 2020–2029.