



From hostile feedback to withdrawal: Emotional exhaustion as a barrier to social media misinformation correction

Simone Moser^{*} , Stephanie Bühner¹ , Diana Kostikova¹ , Jörg Matthes¹ 

Department of Communication, University of Vienna, Vienna, Austria

ARTICLE INFO

Keywords:

Conservation of resources theory
Correction confidence
Emotional exhaustion
Online incivility
Misinformation correction
Social media

ABSTRACT

User-driven misinformation correction is essential for sustaining the integrity of democratic discourse on social media, yet the emotional costs that discourage engagement remain understudied. Drawing on Conservation of Resources (COR) theory, this study examines how the emotional tone of other users' feedback affects correctors' emotional exhaustion and subsequent willingness to engage in future corrections in a social media environment. In a cross-national experiment ($N = 2001$) across Austria, Bulgaria, Denmark, and Poland, utilizing a single-factor, four-level between-subjects design, participants corrected misinformation and received feedback varying in valence and arousal. Results show that angry, high-arousal feedback uniquely depletes emotional resources, producing higher exhaustion than civil disapproval or positive feedback. This indicates that critical feedback does not inherently exhaust correctors, but it is the hostile tone, not the negativity itself, that drives depletion. Emotional exhaustion, in turn, predicts avoidance of future corrective engagement. This exhaustion-driven withdrawal is amplified rather than buffered by correction confidence, suggesting that individual motivation alone is insufficient in hostile environments. These findings highlight emotional exhaustion as a key psychological constraint on citizen-led fact-checking and suggest that platform-level moderation targeting hostile comments is critical to protect the emotional resources required for a resilient democratic debate in digital environments.

1. Introduction

Misinformation has become a defining feature of contemporary social media ecosystems, with direct consequences for democratic discourse. Exposure to false content can distort citizens' political beliefs and systematically bias collective opinion on consequential political issues (Jerit & Zhao, 2020). Despite citizens' unprecedented access to political information, social media simultaneously exposes users to biased and false content that spreads rapidly and at scale (Aïmeur et al., 2023; Bartsch et al., 2024). False news diffuses faster than accurate information (Vosoughi et al., 2018; Weismueller et al., 2024), a dynamic amplified by algorithmically reinforced echo chambers that limit exposure to opposing viewpoints (Lazer et al., 2018) and foster the uncritical circulation of misinformation (Muhammed & Mathew, 2022).

Correcting misinformation presents a particularly pressing challenge (Chan et al., 2017). Although social media platforms have invested in automated moderation and algorithmic detection, such technological approaches face significant constraints. Empirically, automated systems struggle to address the scale and complexity of online information

(Bode, 2020). From an ethical standpoint, concerns persist regarding the concentration of authority in private platforms to define misinformation (Bode, 2020). User-driven misinformation correction thus represents a critical complementary mechanism. Research indicates that a single corrective user intervention can reduce misperceptions not only in the directly addressed individual but also among observing bystanders (Bode & Vraga, 2021). Moreover, user-led corrections are as effective as algorithmic interventions in reducing misperceptions (Bode & Vraga, 2018). Taken together, user-driven correction functions as an essential and ethically grounded complement to platform-based moderation strategies (Bode, 2020).

While existing research has predominantly focused on expert-led corrections and their effectiveness in changing recipients' beliefs (Bautista et al., 2022; Bode & Vraga, 2018; Ecker et al., 2022; Tolz et al., 2025), the factors shaping ordinary users' willingness to correct misinformation remain understudied. Prior research indicates that users anticipate hostile reactions from others in online interactions (Vraga et al., 2015), and that concerns about social consequences discourage corrective engagement (Gurgun et al., 2025). These findings highlight

^{*} Corresponding author. Department of Communication, University of Vienna, Währinger Straße 29 (R. 3.37), Vienna, 1090, Austria.

E-mail address: s.moser@univie.ac.at (S. Moser).

<https://doi.org/10.1016/j.chb.2026.109179>

Received 20 May 2026; Received in revised form 20 August 2026; Accepted 4 September 2026

Available online 14 September 2026

0747-5632/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

the importance of emotional costs associated with engaging in misinformation correction. Initial qualitative evidence suggests that emotional exhaustion constitutes a key factor in corrective behavior (Olley & Ikerodah, 2025). Emotional exhaustion refers to the depletion of emotional resources required to cope with demanding interactions, leaving individuals feeling drained and less capable of sustaining further engagement (Emerson et al., 2023; Maslach & Jackson, 1981), and social media use more broadly has been shown to induce it (Bright et al., 2015; Sheng et al., 2023; Zheng & Ling, 2021). Because correcting misinformation can be socially and emotionally taxing, exhaustion may foster misinformation correction avoidance. Misinformation correction avoidance refers to an individual's future-oriented reluctance to correct misinformation encountered on social media. It is specific to the corrective act, which distinguishes it from broader forms of social media disengagement, or withdrawal from online participation as a whole. From a resource-conservation perspective, correction avoidance reflects motivated withholding of a potentially effortful or conflictual activity, capturing an active inclination to refrain from correcting misinformation in future encounters, rather than a mere absence of intention to act. Yet it remains unknown how specific aspects of corrective interactions – particularly the emotional tone of others' feedback – shape correctors' emotional exhaustion and, in turn, influence their subsequent engagement in misinformation correction.

To address this gap, we conducted a cross-national online experiment across four culturally and politically diverse European countries, in which participants corrected misinformation and then received feedback on their correction that varied in emotional tone. Feedback tone was defined along two dimensions, valence (positive vs. negative) and arousal (high vs. low), yielding four conditions: anger (negative, high-arousal), disapproval (negative, low-arousal), approval (positive, low-arousal), and enthusiasm (positive, high-arousal). Because anger was expected to be the most exhausting, we used it as the reference point, examining whether each non-hostile reaction produced less emotional exhaustion and, in turn, reduced avoidance of future misinformation correction. This comparison reveals whether it is the negativity of feedback, its arousal, or their combination in anger that is especially exhausting.

Our study advances research on misinformation correction online through three key contributions. First, by comparing angry feedback against feedback that shares only its valence, only its arousal, or neither, we show that the combination of negative valence and high arousal in hostile feedback is especially depleting. Thus, we establish emotional costs as a psychologically meaningful and causal consequence of corrective engagement. Second, by extending Conservation of Resources theory (Hobfoll, 2001) to the context of misinformation correction, we introduce a resource-based framework that accounts for both the affective consequences of feedback tone and their downstream behavioral effects. Third, this study shifts the focus of misinformation correction research from recipients to correctors, highlighting their emotional experiences as central to sustaining citizen-led fact-checking. While prior research has primarily examined whether corrections change beliefs, our findings show that continued corrective engagement depends on the emotional climate of online discussions.

Together, these contributions provide a theoretically grounded and empirically robust account of the dynamics underlying misinformation correction. The results indicate that angry, high-arousal feedback, rather than criticism itself, drives emotional exhaustion and discourages continued corrective action. This suggests that citizen-led fact-checking can be sustained even in the presence of disagreement, provided that online discussions remain civil.

2. Theoretical background

2.1. The impact of feedback tone on emotional exhaustion

Psychological barriers are central to users' hesitation to correct

misinformation online, as emotional feedback on social media functions as a powerful affective regulator that shapes subsequent emotions and behaviors (Johnson & Connelly, 2014; Stsiampkouskaya et al., 2021). Russell's (1980) circumplex model represents affect within a two-dimensional space defined by valence (positive vs. negative) and arousal (high vs. low), and both dimensions shape affective experience and behavioral responses (Russell & Barrett, 1999). We argue that emotional social media reactions can be located within this same space, as they vary in whether they signal positivity or negativity and in the intensity of that valenced signal. We use these two dimensions of the model to organize four types of social feedback, deriving one feedback type to represent each quadrant: anger (negative, high arousal), disapproval (negative, low arousal), approval (positive, low arousal), and enthusiasm (positive, high arousal). Anger and enthusiasm correspond closely to affect terms scaled in the original circumplex model. Approval and disapproval are evaluative reactions, expressing calm endorsement or rejection. We conceptualize them as relatively lower-arousal responses because they communicate a positive or negative evaluation without the heightened emotional activation of more intense emotional responses such as anger or enthusiasm. On this basis, we locate approval and disapproval in the low-arousal quadrants, recognizing that they are not themselves affect terms scaled in the original model. This derivation retains the structure of the circumplex model while translating it into discrete forms of feedback that correctors might encounter online.

Emotional exhaustion, as defined above, is a particularly relevant consequence in this context, as it captures the depletion of psychological resources arising from emotionally demanding interactions (Maslach & Jackson, 1981) and has been identified as a key determinant of disengagement from social media environments (Han, 2018; Sheng et al., 2023). Although the construct originates in research on occupational burnout, emotional exhaustion can also be experienced and captured as an acute, momentary state in response to a specific experience. Prior experimental work has adapted established exhaustion measures to a momentary referent and assessed exhaustion as an immediate response within a single session (Sosnowska et al., 2019). In the present study, we adopt this state-based conceptualization and examine emotional exhaustion as the immediate experience of depletion following an emotionally charged online interaction. Emerging in contexts involving intense interpersonal interactions and high-arousal negative affect (Li et al., 2017; Maslach & Jackson, 1981; Shah et al., 2022), exhaustion is likely to arise during misinformation correction on social media, where addressing false beliefs can provoke emotionally charged exchanges.

Conservation of Resources (COR) theory (Hobfoll, 2001) provides a framework for understanding how different types of feedback may have resource-related implications. Valence sets the direction of resource change with negative feedback signaling rejection and threatening social resources such as esteem and standing, whereas positive feedback confers them. Arousal may instead shape the magnitude of that change. Activated feedback could convey the valenced signal more intensely with hostile feedback high in arousal signaling a stronger threat to the corrector's standing, and enthusiastic feedback high in arousal a stronger affirmation. Since COR does not itself address arousal, we propose this amplifying role as a theoretical extension.

Anger, negative in valence and high in arousal, therefore represents the feedback condition with the strongest potential for resource depletion. Drawing on COR (Hobfoll, 2001), anger is linked to direct resource loss. It drains emotional energy and undermines users' sense of competence and perceived control, thereby reducing individuals' capacity to cope with additional resource threats. Anger is thus closely related to emotional exhaustion (Li et al., 2017; Shah et al., 2022).

However, negative valence does not inherently imply high arousal or hostility, as feedback can convey disagreement while remaining emotionally restrained. Disapproval exemplifies this by sharing anger's negative valence while lacking its arousal intensity, thereby signaling disagreement without hostility. While disapproval may still threaten

social resources by undermining feelings of being valued by others (Hobfoll, 2001), research confirms that civil disagreement elicits significantly less emotional discomfort than uncivil disagreement (Chen & Ng, 2017), suggesting that disapproval imposes a smaller resource cost than anger.

Among positive feedback types, approving feedback is relatively low in arousal. It provides acknowledgment and validation that help conserve existing emotional reserves (Hobfoll, 2001). Approving feedback aligns with resources such as acknowledgment of one's accomplishments and affection from others, which Hobfoll (2001) identifies as key conservation targets. Confirmation messages in interpersonal contexts have been linked to adaptive coping (Dailey, 2023), and positive affectivity more broadly has been associated with reduced exhaustion (Szczygiel & Bazińska, 2013).

Enthusiastic feedback combines positive valence with high arousal. Within COR's framework (Hobfoll, 2001), such feedback may represent a potential source of resource gain, fostering positive motivational resources that expand emotional reserves and may initiate gain spirals in which resource acquisition facilitates further gain (Hobfoll, 2001). This aligns with broaden-and-build theory (Fredrickson, 2001), which states that positive emotions expand individuals' cognitive and behavioral repertoires and thereby generate new personal resources. In social media contexts, enthusiastic responses such as encouraging comments have been shown to evoke excitement and sustain engagement (Stsiampkouskaya et al., 2021).

Taken together, COR theory suggests that resource loss is disproportionately more salient than resource gain, and that initial losses can accelerate into further loss spirals (Hobfoll, 2001), positioning angry feedback as the most severe resource threat. Our central aim is therefore to test whether anger is distinct in its emotional cost, such that its combination of high arousal and negative valence is more depleting than feedback that shares only its valence, only its arousal, or neither. Comparing anger to disapproval holds valence constant and varies arousal, indicating whether arousal intensifies exhaustion within negative feedback. Comparing anger to enthusiasm holds arousal constant and varies valence, allowing us to examine whether an intense emotional reaction is sufficient to produce greater exhaustion or whether negative valence is also required. Comparing anger to approval reflects the combined difference on both dimensions. Together, these comparisons clarify whether anger's distinct cost reflects its negativity, its high arousal, or their combination. Accordingly, we hypothesize:

H1a. Angry feedback will lead to higher emotional exhaustion than disapproving feedback.

H1b. Angry feedback will lead to higher emotional exhaustion than approving feedback.

H1c. Angry feedback will lead to higher emotional exhaustion than enthusiastic feedback.

2.2. Emotional exhaustion as a barrier to misinformation correction

In addition to its development, the behavioral consequences of emotional exhaustion can also be explained with COR theory (Hobfoll, 2001). When stressors deplete emotional resources, individuals prioritize conservation of remaining reserves over further investment to prevent additional loss. In this state, effortful and socially risky behaviors are likely to be avoided, as they require resources that exhausted individuals cannot afford. This pattern manifests as withdrawal from demanding social media use, with emotional exhaustion shown to reduce knowledge-sharing (Zhao, 2024) and increase disengagement from online interactions more broadly (Maier et al., 2012).

Misinformation correction on social media represents a particularly demanding interaction, as it combines active resource investment with direct exposure to others' reactions on contested topics. As an inherently confrontational act that risks provoking conflict (Woods & Ruscher,

2021), it's particularly susceptible to the depletion dynamics described above. Qualitative evidence confirms that exhaustion functions as a central barrier to corrective engagement (Olley & Ikerodah, 2025). Whether emotional exhaustion causally reduces users' willingness to correct misinformation in the future, however, remains underexplored. Accordingly, we hypothesize:

H2. Emotional exhaustion leads to higher misinformation correction avoidance in the future.

2.3. The mediating role of emotional exhaustion

Together, the preceding arguments suggest that feedback tone influences misinformation correction avoidance via emotional exhaustion as an underlying mechanism. Angry feedback depletes emotional resources, and this depletion in turn motivates individuals to protect their remaining resources by disengaging from future misinformation correction. Crucially, this effect is expected to be specific to angry feedback, as disapproval, approval, and enthusiasm each impose lower resource costs or actively replenish resources, and are therefore less likely to trigger the resource protection response that drives avoidance. This indirect process is consistent with COR theory (Hobfoll, 2001), which explicitly links resource loss to downstream behavioral consequences through the mediating experience of strain. Yet while the emotional consequences of social media interactions have been studied in general contexts (e.g., Sheng et al., 2023), experimental evidence is lacking on whether feedback tone shapes corrective behavior through emotional exhaustion as the underlying mechanism. Accordingly, we hypothesize:

H3. The effect of angry feedback on misinformation correction avoidance, relative to (a) disapproving feedback, (b) approving feedback, and (c) enthusiastic feedback is mediated by emotional exhaustion.

2.4. Correction confidence as an internal resource for engagement

While emotional exhaustion fosters avoidance tendencies, individuals may differ in how strongly depletion translates into withdrawal. We argue that correction confidence functions as a key moderating resource in this process. Defined as the belief in the accuracy and effectiveness of one's corrective action, it reflects the broader concept of self-efficacy, namely the belief in one's ability to successfully perform a given behavior (Bandura, 1997).

Two complementary theoretical accounts support this role. COR theory (Hobfoll, 2001) argues that those with greater internal resources are better buffered against resource loss, as a stronger resource reservoir interrupts the path from strain to behavioral consequences. Self-efficacy constitutes precisely such a resource: according to Social Cognitive Theory (Bandura, 1997), individuals with high self-efficacy beliefs are less affected by stressful encounters because they appraise demanding situations as manageable rather than threatening, reducing the behavioral consequences of strain. Empirically, self-efficacy is negatively associated with work and social withdrawal (Zhong et al., 2023), and positively linked to engagement behaviors (Hong et al., 2021). In digital environments specifically, higher self-efficacy predicts stronger social media engagement (Saidah & Hariyadi, 2023).

Applied to misinformation correction, users who are confident in the accuracy and effectiveness of their corrections may be less inclined to avoid future corrective engagement. Despite the established link between self-efficacy and engagement behaviors across digital contexts, it remains unknown whether this extends to the specific case of misinformation correction, where social and emotional costs are particularly pronounced. Accordingly, we hypothesize:

H4a. Higher correction confidence is associated with lower misinformation correction avoidance.

Beyond its direct effect on avoidance, correction confidence may also buffer the impact of emotional exhaustion on corrective withdrawal. Self-efficacy has been shown to reduce the impact of perceived stress on exhaustion-related outcomes (Prati et al., 2010), suggesting that internal efficacy resources might interrupt the path from strain to disengagement. Thus, confident correctors may appraise the emotional costs of others' feedback as more manageable, weakening the extent to which exhaustion translates into avoidance. Yet whether efficacy beliefs attenuate exhaustion-driven withdrawal remains empirically unexamined, leaving open the question of which internal resources determine who continues to correct misinformation despite its emotional costs. Accordingly, we propose:

H4b. Correction confidence moderates the relationship between emotional exhaustion and misinformation correction avoidance, such that higher correction confidence attenuates the positive effect of emotional exhaustion on avoidance.

Fig. 1 illustrates the theoretical model underlying this study.

2.5. Cross-national variation and generalizability

We examined four European countries (Austria, Bulgaria, Denmark, and Poland) selected to provide variation along contextual dimensions relevant to misinformation and emotional dynamics. Their information environments differ in disinformation resilience, from Denmark among the most resilient EU states to Bulgaria among the lowest, with Austria and Poland intermediate (Bleyer-Simon et al., 2025), and in the independence of media institutions, from fully independent in Austria and Denmark to state-captured public media in Bulgaria and a transitioning system in Poland (Bleyer-Simon et al., 2025). They are further distinguished by their levels of political polarization, censorship (Coppedge et al., 2025; OSF Fig. A1), and media freedom (Reporters Without Borders, 2025). These structural differences extend to the character of national disinformation itself, shaped in Austria by longstanding racist and antisemitic narratives, in Poland by regional geopolitical pressures, and in Bulgaria by institutional distrust and political instability (Margova & Dobрева, 2023; Schäfer, 2023; Zadroga, 2025).

Culturally, they span more individualistic (Austria, Denmark) and more collectivistic (Bulgaria, Poland) orientations and differ in uncertainty avoidance, achievement orientation, power distance, and indulgence (Minkov & Kaasa, 2022; The Culture Factor Group, 2023; see OSF Fig. A2), dimensions linked to emotional norms (Eid & Diener, 2001). In more collectivistic contexts, emotions are experienced as relational phenomena tied to social standing and worth, whereas in more individualistic contexts they are treated more as intrapersonal states (Mesquita, 2001). Norms governing negative emotion likewise differ: across many countries, including Denmark and Poland, individualistic

cultures permit greater expression of negative emotion toward in-groups, whereas collectivistic cultures tend to suppress such expression to preserve group harmony (Matsumoto et al., 2008). At the same time, adherence to emotion norms is greater in individualistic cultures, where deviating from others' emotional experience more strongly predicts reduced well-being (Vishkin et al., 2023). These differences could shape how recipients interpret the social meaning and resource implications of corrective feedback, such that the same feedback may translate into exhaustion differently across contexts. However, the existing literature does not provide a sufficiently clear basis for predicting the direction of such cross-national differences. Nonetheless, the online environments share important features across national contexts. Negativity is a consistent driver of engagement with political content across a wide range of European countries, including Austria, Denmark, and Poland (Bene et al., 2022), indicating that hostile feedback is a common feature of the settings correctors encounter across various European contexts.

Taken together, this combination of contextual variation and shared features of the online environment makes a cross-national design particularly informative for assessing whether the proposed mechanism generalizes across settings. We consequently treat cross-national scope as a test of generalizability, where consistent support for the hypothesized process across these contexts would indicate a mechanism robust to cultural and institutional variation.

3. Method

The study was conducted by Dynata in Austria, Bulgaria, Denmark, and Poland. Following ethical approval granted by the Institutional Review Board of the University of Vienna, data collection took place between December 22, 2025 and January 12, 2026. All supplementary materials, including item wordings in all languages, the dataset, descriptive statistics, analysis scripts, and outputs, are publicly available on OSF (<https://osf.io/5hdqg>).

3.1. Participants

In each country, quota-based sampling ensured representativeness across gender, age, and education. Participants were eligible if they (a) were citizens and residents of Austria, Bulgaria, Denmark, or Poland, (b) were between 18 and 69 years old (c) had used social media (e.g., WhatsApp, Twitter/X), within the two months prior, and (d) provided informed consent. Following data collection, participants were excluded if they (a) failed both attention checks ($n_{\text{Austria}} = 98$, $n_{\text{Bulgaria}} = 121$, $n_{\text{Denmark}} = 122$, $n_{\text{Poland}} = 75$), (b) completed the survey in less than one-third of the median completion time for their country sample (speeders; Greszki et al., 2014; $n_{\text{Austria}} = 5$, $n_{\text{Bulgaria}} = 2$, $n_{\text{Denmark}} = 4$, $n_{\text{Poland}} = 1$),

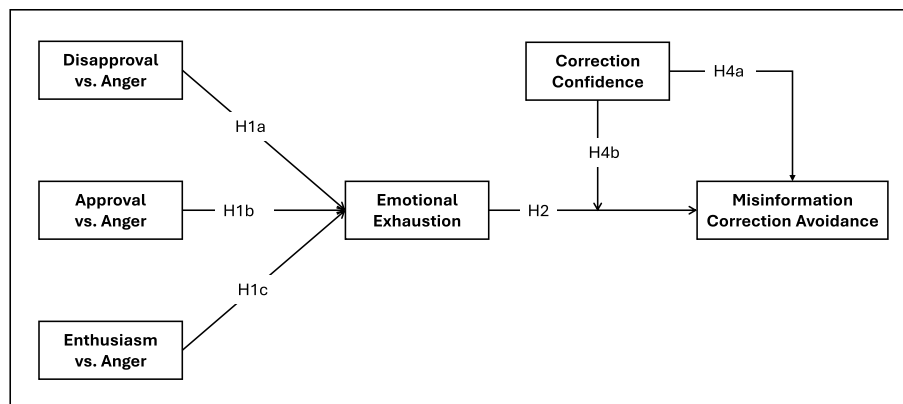


Fig. 1. Theoretical Model.

Note. Control variables, measurement errors, and mediation paths were omitted for clarity. H = hypothesis.

or (c) identified their gender as other than male or female, as the resulting subgroup was too small for meaningful analysis ($n_{\text{Austria}} = 0$, $n_{\text{Bulgaria}} = 0$, $n_{\text{Denmark}} = 0$, $n_{\text{Poland}} = 1$). The final sample comprised $N = 2001$ participants ($n_{\text{Austria}} = 494$, $n_{\text{Bulgaria}} = 501$, $n_{\text{Denmark}} = 505$, $n_{\text{Poland}} = 501$). Demographic details are presented in OSF Table A1.

Cross-country comparisons showed no notable differences in control variables. Although some differences were statistically significant, effect sizes were negligible by conventional standards and unlikely to meaningfully influence the results (OSF Tables A2–A3).

3.2. Procedure and experimental manipulation

All materials were originally developed in English and translated into German, Bulgarian, Danish, and Polish using AI-based translation tools (DeepL, ChatGPT). Native-speaking researchers subsequently reviewed all translations to verify linguistic accuracy and conceptual equivalence across language versions. Item and stimuli wordings in all languages (Tables A4–A23 and Fig. A3–A52), and item-level statistics (Tables A24–A47) are openly accessible on OSF.

3.2.1. Procedure

The experiment was conducted online and employed a single-factor, four-level between-subjects design. Following informed consent, participants completed introductory questions on demographics and social media use. They were then randomly assigned to one of four feedback conditions (anger, disapproval, approval, or enthusiasm; see OSF Table A48) and completed two consecutive experimental rounds with identical procedures.

Each round began with participants being informed that the post they were about to see contained misinformation, after which they were presented with a distinct fabricated Facebook-style post claiming that AI would displace human workers (Fig. 2 shows the English version of one of the two posts; OSF Fig. A3–A12 provide all language versions). This topic was chosen because it is affectively relevant and comparable across the four countries while avoiding ideologically charged topics. Each post

presented specific fabricated claims (a secret government plan to replace most jobs with AI within a few years, attributed to leaked documents and anonymous insiders) so that the misinformation post was clearly false regardless of the broader debate about AI and employment. Participants were subsequently instructed to correct that misinformation post by selecting one of nine predefined correction statements (e.g., “Evidence from experts does not confirm that claim”). All options were developed around the same core corrective message that the claim was unsupported by reliable evidence, data, or expert sources. They were designed to be equivalent in factual content, tone, and directness, while varying in phrasing. Participants were thus offered a choice among equivalent options so that the correction would be a more active, self-determined act. The selection did not affect subsequent stimuli or measures.

After the correction task, participants were presented with a condition-specific simulated comment section displaying other users’ reactions to their correction (Fig. 3 shows one English-language example per condition; OSF Figures A13–A52 show the full set, including both rounds and all language versions). Round-specific measures were completed after each round. After the second round, participants completed a final set of general measures on misinformation engagement and political attitudes, followed by a thorough debriefing.

3.2.2. Experimental stimuli and manipulation

The simulated comment sections constituted the experimental manipulation. As part of the experimental cover story, participants were told that the comments reflected response patterns characteristic for their fellow citizens, based on a nationally conducted pretest. This cover story was intended to enhance the perceived authenticity and credibility of the feedback. In reality, the comment sets were constructed to convey the targeted emotional tone of their respective experimental condition (anger, disapproval, approval, or enthusiasm; see Fig. 3). Each comment section comprised seven comments: five conveying the emotional tone of the assigned condition, and two neutral filler comments included to enhance perceived authenticity. To ensure comparability across conditions, user identities and structural features of the comment sections were held constant. Prior to the main study, all comment sets underwent pretesting with an independent sample ($N = 419$) to confirm stimulus validity. Only sets that conveyed and elicited the intended emotion, were retained.

3.3. Measures

Descriptive statistics for all path model variables by condition and country are presented in Table 1 (for covariates, see OSF Table A49; for reliability coefficients, see OSF Table A50). Confirmatory factor analysis (CFA) indicated good-to-acceptable fit across most measurement models (OSF Tables A51–A61), with the exception of correction confidence in Austria and Poland and emotional exhaustion in Poland. Tests of measurement invariance (OSF Tables A62–A66) supported metric and configural invariance across most scales, with the exception of correction confidence. Scalar and strict invariance were not fully supported across scales, however, partial invariance is considered sufficient to support valid cross-national comparisons (Davidov et al., 2014). ANOVAs indicated significant cross-country differences in emotional exhaustion and misinformation correction avoidance, but not in correction confidence (OSF Table A2). Effect sizes were negligible ($\eta^2 < .01$ –.02), suggesting these differences have little practical significance.

Because misinformation correction avoidance, our final outcome, was assessed only once after the second experimental round, the Round 2 measures of emotional exhaustion and correction confidence are the assessments temporally closest to this outcome, and we retained them for the primary model. For consistency across round-specific measures, Round 2 assessments of perceived misinformation accuracy and fact-checking intentions were used as well. In practice, all round-specific measures were largely stable across rounds (see Section 4.5). Re-estimating the model using mean scores across both rounds yielded a

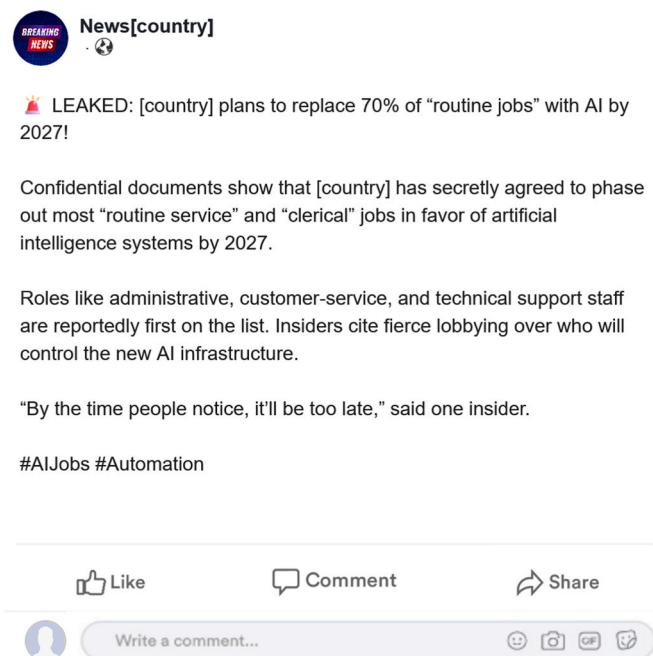


Fig. 2. Example of the fabricated misinformation posts.

Note. The post is shown in English language for illustration; in the experiment, posts were presented in the respective country's language, with content and structural features held constant across countries. [Country] indicates the respective country.

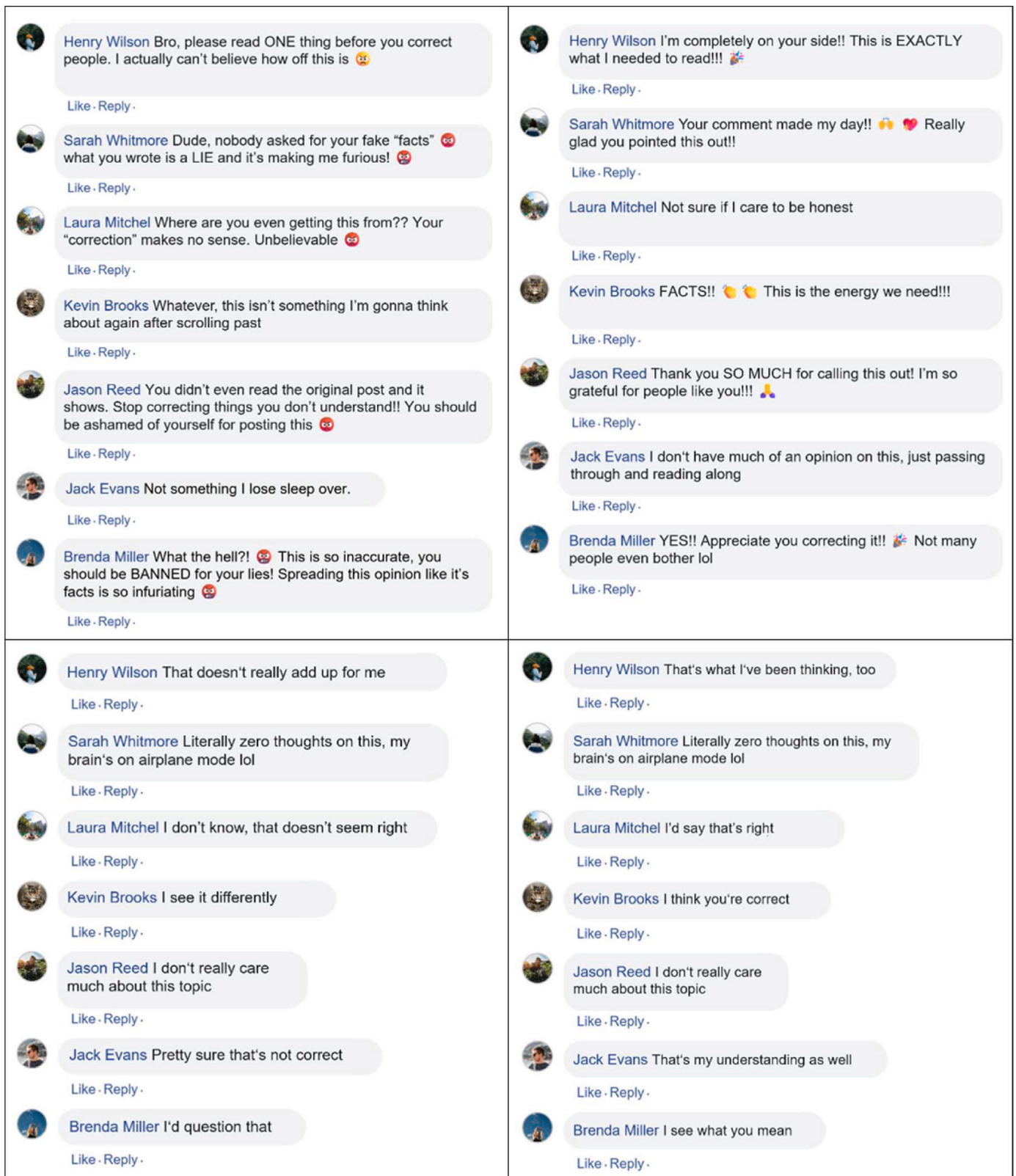


Fig. 3. Examples of the condition-specific simulated comment sections.

Note. Angry (top left), enthusiastic (top right), disapproving (bottom left), and approving (bottom right) feedback in English language for illustration; in the experiment, comments were presented in the respective country's language, with content and structural features held constant across countries. Profile names were adapted to reflect common names in each country.

Table 1

Means and standard deviations of all variables per condition and country.

Experimental Condition	Variable	Austria	Bulgaria	Denmark	Poland
		<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>	<i>M (SD)</i>
Enthusiasm	Emotional Exhaustion	1.88 (1.08)	2.07 (1.22)	2.12 (1.24)	1.93 (1.22)
	Misinformation Correction Avoidance	2.73 (.89)	2.61 (1.12)	2.82 (1.00)	2.74 (1.23)
	Correction Confidence	3.56 (.91)	3.71 (.95)	3.67 (.92)	3.83 (.96)
Approval	Emotional Exhaustion	1.90 (1.16)	2.16 (1.18)	2.28 (1.24)	2.06 (1.13)
	Misinformation Correction Avoidance	2.85 (1.06)	2.74 (1.19)	3.01 (.84)	3.01 (1.03)
	Correction Confidence	3.43 (1.05)	3.62 (1.07)	3.51 (.85)	3.72 (.92)
Disapproval	Emotional Exhaustion	1.68 (.91)	1.94 (1.08)	2.03 (1.17)	1.93 (1.17)
	Misinformation Correction Avoidance	2.70 (1.14)	2.49 (1.00)	3.09 (1.09)	3.02 (1.15)
	Correction Confidence	3.54 (1.02)	3.65 (1.04)	3.50 (.99)	3.47 (1.04)
Anger	Emotional Exhaustion	2.26 (1.25)	2.32 (1.26)	2.56 (1.18)	2.19 (1.19)
	Misinformation Correction Avoidance	2.92 (1.12)	2.54 (1.21)	2.97 (.96)	2.95 (1.12)
	Correction Confidence	3.40 (.95)	3.53 (1.20)	3.59 (.86)	3.57 (1.05)

substantively similar central pattern (see Section 4.5). The choice of round therefore does not substantively affect the results.

3.3.1. Emotional exhaustion

Emotional exhaustion was measured with four items adapted from Maslach et al. (1997) and Han (2018), assessing how participants felt after reading the comment section on a five-point Likert scale (1 = *not at all*, 5 = *extremely*; e.g., “I feel emotionally drained”, “These comments put too much stress on me”). Whereas these sources assess exhaustion as a general or chronic state, we anchored the items to the present moment (“how you feel at this moment, when thinking about the comments”), measuring exhaustion as an immediate response to feedback. This momentary adaptation follows prior experimental work that likewise adjusted established exhaustion items to a momentary referent within a single session (Sosnowska et al., 2019). The scale demonstrated good to excellent reliability in each country sample (Cronbach's α : $\alpha_{\text{Austria}} = .92$, $\alpha_{\text{Bulgaria}} = .92$, $\alpha_{\text{Denmark}} = .95$, $\alpha_{\text{Poland}} = .94$; McDonald's ω : $\omega_{\text{Austria}} = .93$, $\omega_{\text{Bulgaria}} = .92$, $\omega_{\text{Denmark}} = .95$, $\omega_{\text{Poland}} = .94$).

3.3.2. Misinformation correction avoidance

Misinformation correction avoidance was captured with four items reflecting reluctance to engage in future correction, rated on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*; e.g., “I think correcting misinformation on this topic is not worth my time.”; “I don't want to risk arguments, so I'll stay silent about false claims on this topic.”). The items assess the intention to refrain from future correction through its characteristic barriers, including perceived effort or futility of correcting, avoidance of conflict, and the attribution of responsibility to others (Bode & Vraga, 2021; Gurgun et al., 2025). Internal consistency was good across all four country samples ($\alpha_{\text{Austria}} = .83$, $\alpha_{\text{Bulgaria}} = .86$, $\alpha_{\text{Denmark}} = .81$, $\alpha_{\text{Poland}} = .87$; $\omega_{\text{Austria}} = .83$, $\omega_{\text{Bulgaria}} = .86$, $\omega_{\text{Denmark}} = .81$, $\omega_{\text{Poland}} = .87$).

3.3.3. Correction confidence

Correction confidence was assessed with four items measuring participants' post-feedback appraisal of the perceived accuracy and effectiveness of their correction, rated on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*; e.g., “I chose a decent response option to correct the misinformation.”, “My correction effectively corrected the misinformation.”). Reliability coefficients indicated good internal consistency across countries ($\alpha_{\text{Austria}} = .83$, $\alpha_{\text{Bulgaria}} = .88$, $\alpha_{\text{Denmark}} = .83$, $\alpha_{\text{Poland}} = .89$; $\omega_{\text{Austria}} = .83$, $\omega_{\text{Bulgaria}} = .89$, $\omega_{\text{Denmark}} = .84$, $\omega_{\text{Poland}} = .89$).

3.3.4. Covariates

Gender, age, and education were included as covariates to control for potential confounding effects. Gender was operationalized as a binary variable (male/female), age was self-reported numerically, and education was captured via country-specific categories reflecting participants'

highest completed qualification, subsequently dichotomized into university-educated and non-university-educated groups for analysis.

Political orientation, social media literacy, fact-checking intentions, and perceived accuracy of misinformation were included as additional covariates. Political orientation was assessed on a five-point left–right scale. Social media literacy was measured with seven items adapted from Koc and Barut (2016), capturing participants' self-reported knowledge and skills for navigating social media on a five-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*; e.g., “I can assess social media content in terms of credibility, reliability, objectivity and currency”). Fact-checking intentions were assessed with seven items adapted from Flanagan and Metzger (2000), reflecting the frequency of verifying online information on a five-point Likert Scale (1 = *never*, 5 = *all the time*; e.g., “I seek out other sources to validate the information.”). Internal consistency was good for social media literacy ($\alpha_{\text{Austria}} = .85$, $\alpha_{\text{Bulgaria}} = .87$, $\alpha_{\text{Denmark}} = .88$, $\alpha_{\text{Poland}} = .88$; $\omega_{\text{Austria}} = .85$, $\omega_{\text{Bulgaria}} = .87$, $\omega_{\text{Denmark}} = .88$, $\omega_{\text{Poland}} = .88$) and excellent for fact-checking intentions ($\alpha_{\text{Austria}} = .92$, $\alpha_{\text{Bulgaria}} = .91$, $\alpha_{\text{Denmark}} = .91$, $\alpha_{\text{Poland}} = .92$; $\omega_{\text{Austria}} = .92$, $\omega_{\text{Bulgaria}} = .91$, $\omega_{\text{Denmark}} = .91$, $\omega_{\text{Poland}} = .92$). Perceived accuracy of misinformation was measured with a single item adapted from Martel et al. (2020), asking participants to rate the accuracy of the misinformation post they had just read (1 = *not at all accurate*, 4 = *very accurate*).

3.4. Statistical analysis

3.4.1. Manipulation checks

To verify the effectiveness of the experimental manipulation, immediately after each feedback exposure, participants rated how strongly each of four affective tones – anger, disapproval, approval, and enthusiasm – was expressed in the comments (1 = *not at all*, 5 = *extremely*). To account for cross-country differences in response styles, ratings were z-standardized within each country. Planned within-condition contrasts compared ratings of the primed affective tone against each non-primed alternative using one-sided paired *t*-tests. Across almost all countries and both rounds, the intended affective tone significantly received the highest mean rating in its corresponding condition, confirming successful manipulation. Consistent with theoretical expectations, the enthusiasm condition additionally elicited increased perceptions of approval, and the anger condition increased perceptions of disapproval (see OSF Tables A67–A74 and OSF Figures A53–A54). An exception was observed in Poland during Round 1, where the approval condition showed no significant differentiation from disapproval ($p = .966$) or anger ($p = .383$), and the enthusiasm condition did not differ significantly from disapproval ($p = .072$). Notably, however, the intended affective tone still received the highest mean rating in both conditions.

3.4.2. Statistical modeling

All analyses were conducted in R, estimating a single path model simultaneously across all four country samples. A path model was specified with experimental conditions dummy-coded using the anger condition as the reference group, since anger was theoretically predicted to induce the highest levels of emotional exhaustion, allowing each contrast to directly quantify the reduction in exhaustion associated with non-hostile feedback types. Emotional exhaustion was modeled as a mediator between feedback condition and misinformation correction avoidance, with correction confidence included as a moderator of the exhaustion-to-avoidance path. Missing data were handled through full-information maximum likelihood estimation (FIML; Lee & Shi, 2021). The stability of indirect effects was evaluated using 5000 bootstrap samples. To examine cross-country variation, a freely estimated model was compared against a series of multi-group models in which individual paths were sequentially constrained to equality across countries, following the approach described by Hayes et al. (2013). Gender, age, education, political orientation, social media literacy, perceived accuracy of misinformation, and fact-checking intentions were included as covariates throughout.

In response to the theoretically proposed valence–arousal structure of the feedback conditions, we additionally estimated a complementary factorial parameterization in which the four conditions were coded as 2 (valence: negative vs. positive) $\times$ 2 (arousal: high vs. low). This model included the valence $\times$ arousal interaction and follow-up simple effects to examine whether the effect of either dimension depended on the level of the other. Because valence and arousal were not directly rated as separate dimensions, these analyses are interpreted as a decomposition of the theoretically intended characteristics of the four feedback conditions rather than as evidence of independently validated dimensional manipulations. The complementary factorial model retained the same covariates and model specification as the primary analysis, with only the coding of the experimental conditions changed.

4. Results

Zero-order correlations by country are provided on OSF Tables A75–A78. Table 2 summarizes the main path analysis results for all countries. Mediation results are detailed in Table 3. Full results, including covariate estimates, are available in OSF Tables A79–A80. Fig. 4 illustrates the model structure and key findings. Multi-group comparisons (OSF Tables A81–A84) revealed that most individual path estimates did not differ significantly across countries, supporting the cross-national stability of the main findings. An exception was the interaction between emotional exhaustion and correction confidence, which varied significantly across countries ($\Delta\chi^2(3) = 10.31, p = .005$).

Table 2
Path analysis results using anger as reference group.

	Austria (n = 494)		Bulgaria (n = 501)		Denmark (n = 505)		Poland (n = 501)	
	B (SE), β	p	B (SE), β	p	B (SE), β	p	B (SE), β	p
<i>Emotional Exhaustion Round 2</i>								
Enthusiasm (vs. Anger)	-.40 (.14), -.15	.005	-.30 (.16), -.11	.051	-.47 (.13), -.16	< .001	-.33 (.15), -.12	.023
Approval (vs. Anger)	-.37 (.14), -.14	.011	-.20 (.15), -.07	.183	-.30 (.14), -.11	.026	-.27 (.14), -.10	.049
Disapproval (vs. Anger)	-.57 (.12), -.22	< .001	-.37 (.14), -.13	.010	-.58 (.13), -.20	< .001	-.36 (.14), -.13	.010
R ²	.193		.077		.217		.148	
<i>Misinformation Correction Avoidance</i>								
Emotional Exhaustion	.14 (.05), .15	.003	.19 (.05), .20	< .001	.16 (.04), .20	< .001	.16 (.05), .17	< .001
Correction Confidence	-.53 (.11), -.50	< .001	-.22 (.12), -.20	.062	-.45 (.11), -.42	< .001	-.49 (.10), -.43	< .001
Em. Exh. x Corr. Con.	.20 (.05), .37	< .001	.01 (.05), .02	.864	.16 (.05), .33	.001	.11 (.05), .21	.016
Enthusiasm (vs. Anger)	-.13 (.13), -.05	.303	.14 (.14), .05	.322	-.05 (.11), -.02	.661	-.14 (.14), -.05	.321
Approval (vs. Anger)	-.06 (.13), -.03	.630	.25 (.14), .10	.079	.08 (.10), .04	.434	.09 (.13), .03	.467
Disapproval (vs. Anger)	-.10 (.14), -.04	.454	.06 (.14), .02	.653	.14 (.12), .06	.260	.03 (.14), .01	.850
R ²	.162		.104		.188		.163	

Note. Significant coefficients in bold.

4.1. Effects of feedback tone on emotional exhaustion

Consistent with H1a across all four countries, disapproving feedback significantly predicted lower emotional exhaustion relative to angry feedback in Austria ($B = -.57, SE = .12, \beta = -.22, p < .001$), Bulgaria ($B = -.37, SE = .14, \beta = -.13, p = .010$), Denmark ($B = -.58, SE = .13, \beta = -.20, p < .001$), and Poland ($B = -.36, SE = .14, \beta = -.13, p = .010$).

Supporting H1b, approving feedback significantly predicted lower emotional exhaustion relative to angry feedback in Austria ($B = -.37, SE = .14, \beta = -.14, p = .011$), Denmark ($B = -.30, SE = .14, \beta = -.11, p = .026$), and Poland ($B = -.27, SE = .14, \beta = -.10, p = .049$). This effect was not significant in Bulgaria ($B = -.20, SE = .15, \beta = -.07, p = .183$), contradicting H1b.

Consistent with H1c, enthusiastic feedback significantly predicted lower emotional exhaustion relative to angry feedback in Austria ($B = -.40, SE = .14, \beta = -.15, p = .005$), Denmark ($B = -.47, SE = .13, \beta = -.16, p < .001$), and Poland ($B = -.33, SE = .15, \beta = -.12, p = .023$). This effect was not significant in Bulgaria ($B = -.30, SE = .16, \beta = -.11, p = .051$), failing to support H1c.

To complement the planned feedback-type contrasts, we additionally reparameterized the four conditions as a 2 (valence: negative vs. positive) $\times$ 2 (arousal: high vs. low) factorial structure. A significant valence $\times$ arousal interaction emerged in all four countries (Austria: $B = -.60, SE = .18, \beta = -.13, p = .001$; Bulgaria: $B = -.48, SE = .21, \beta = -.10, p = .020$; Denmark: $B = -.74, SE = .19, \beta = -.15, p < .001$; Poland: $B = -.42, SE = .20, \beta = -.09, p = .032$; see OSF Table A85 for detailed results). Follow-up simple-effect analyses (OSF Table A86) showed that, within negative feedback, high-arousal feedback produced significantly greater emotional exhaustion than low-arousal feedback in all four countries (anger vs. disapproval), whereas arousal did not significantly differentiate the two positive feedback conditions (enthusiasm vs. approval). Likewise, within high-arousal feedback, positive feedback was associated with significantly lower exhaustion than negative feedback in all four countries (enthusiasm vs. anger), whereas valence did not significantly differentiate the two low-arousal conditions (approval vs. disapproval). Thus, the factorial decomposition indicates that the effect of arousal depends on valence and vice versa, with the negative/high-arousal combination represented by anger producing the highest exhaustion.

4.2. Effect of emotional exhaustion on misinformation correction avoidance

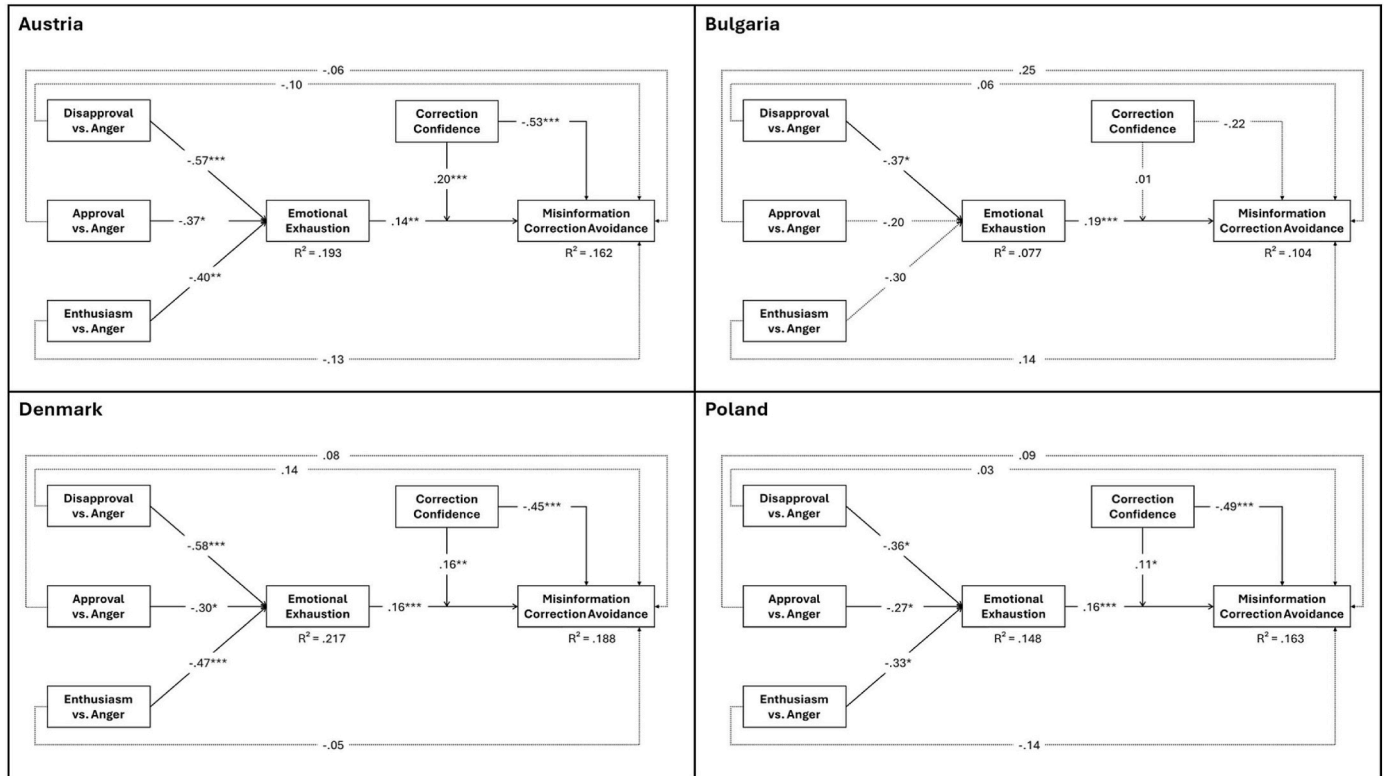
Emotional exhaustion consistently and significantly predicted higher misinformation correction avoidance across all four country samples (Austria: $B = .14, SE = .05, \beta = .15, p = .003$; Bulgaria: $B = .19, SE = .05, \beta = .20, p < .001$; Denmark: $B = .16, SE = .04, \beta = .20, p < .001$; Poland: $B = .16, SE = .05, \beta = .17, p < .001$), supporting H2 in all countries.

Table 3

Indirect and total effects of emotional exhaustion round 2 as mediator and anger as reference group.

	Austria (n = 494)		Bulgaria (n = 501)		Denmark (n = 505)		Poland (n = 501)	
	95% CI [LL, UL]	p	95% CI [LL, UL]	p	95% CI [LL, UL]	p	95% CI [LL, UL]	p
<i>Misinformation Correction Avoidance – Indirect Effects</i>								
Enthusiasm (vs. Anger)	[-.119, -.009]	.057	[-.129, .002]	.078	[-.145, -.025]	.015	[-.122, -.007]	.072
Approval (vs. Anger)	[-.110, -.007]	.061	[-.102, .019]	.210	[-.109, -.005]	.063	[-.101, .001]	.095
Disapproval (vs. Anger)	[-.148, -.024]	.014	[-.146, -.013]	.033	[-.167, -.038]	.004	[-.121, -.011]	.044
<i>Misinformation Correction Avoidance – Total Effects</i>								
Enthusiasm (vs. Anger)	[-.437, .060]	.134	[-.199, .354]	.576	[-.339, .090]	.259	[-.480, .080]	.173
Approval (vs. Anger)	[-.364, .138]	.388	[-.073, .504]	.141	[-.170, .236]	.764	[-.204, .302]	.707
Disapproval (vs. Anger)	[-.442, .086]	.178	[-.279, .262]	.937	[-.192, .282]	.714	[-.300, .234]	.815

Note. Significant coefficients in bold.

**Fig. 4.** Overview Plot for Path Analysis Results.

Note. Additional control variables, the mediation paths (H3), and measurement errors were omitted for clarity. Solid lines indicate significant paths, dashed lines indicate non-significant paths. * $p < .05$, ** $p < .01$, *** $p < .001$.

4.3. Mediating effect of emotional exhaustion

Disapproving feedback (versus anger) produced a significant indirect effect across all four countries (Austria: 95% CI [-.148, -.024], $p = .014$; Bulgaria: 95% CI [-.146, -.013], $p = .033$; Denmark: 95% CI [-.167, -.038], $p = .004$; Poland: 95% CI [-.121, -.011], $p = .044$). The indirect effect of enthusiastic feedback (vs. anger) was significant only in Denmark (95% CI [-.145, -.025], $p = .015$), but remained non-significant in Austria, Bulgaria, and Poland. Approving feedback (vs. anger) did not produce a significant indirect effect in any country (Austria: 95% CI [-.110, -.007], $p = .061$; Bulgaria: 95% CI [-.102, .019], $p = .210$; Denmark: 95% CI [-.109, -.005], $p = .063$; Poland: 95% CI [-.101, .001], $p = .095$). Total effects were non-significant across all conditions and countries. H3 was therefore only partially supported, with robust evidence emerging only for H3a.

4.4. Direct and moderating effects of correction confidence

Consistent with H4a, correction confidence negatively predicted misinformation correction avoidance in Austria ($B = -.53$, $SE = .11$, $\beta = -.50$, $p < .001$), Denmark ($B = -.45$, $SE = .11$, $\beta = -.42$, $p < .001$), and Poland ($B = -.49$, $SE = .10$, $\beta = -.43$, $p < .001$). This effect was not significant in Bulgaria ($B = -.22$, $SE = .12$, $\beta = -.20$, $p = .062$), contradicting H4a.

Contrary to H4b, the interaction between emotional exhaustion and correction confidence was significantly positive in Austria ($B = .20$, $SE = .05$, $\beta = .37$, $p < .001$), Denmark ($B = .16$, $SE = .05$, $\beta = .33$, $p = .001$), and Poland ($B = .11$, $SE = .05$, $\beta = .21$, $p = .016$), indicating that higher correction confidence amplified rather than attenuated the effect of emotional exhaustion on avoidance. The interaction was non-significant in Bulgaria ($B = .01$, $SE = .05$, $\beta = .02$, $p = .864$), also failing to support H4b.

Because correction confidence was assessed after exposure to the experimental feedback, we additionally examined whether feedback

condition predicted Round 2 correction confidence. No significant condition effects emerged in Austria, Bulgaria, or Denmark (all $ps \geq .416$, $\eta^2 \leq .006$). In Poland, the omnibus effect was small, $F(3, 497) = 3.22$, $p = .023$, $\eta^2 = .019$, but did not remain significant after Holm correction across the four country-level tests ($p = .091$; see OSF Table A87 for details). Follow-up comparisons (OSF Table A88) indicated higher confidence following enthusiastic than disapproving feedback in Poland, whereas no other pairwise differences were significant. Overall, feedback condition therefore showed no robust or consistent association with correction confidence across countries.

4.5. Robustness check

To assess whether the findings depended on the use of Round 2 measures in the main analysis, we conducted two complementary robustness checks. First, we compared all measures assessed in both experimental rounds within each country and feedback condition (OSF Tables A89–A92). Overall, the measures were largely stable across rounds. For emotional exhaustion, most Round 1–Round 2 differences were non-significant after Holm correction; significant decreases emerged only in the enthusiasm and approval conditions in Denmark, whereas exhaustion in the anger condition did not differ significantly between rounds in any country (OSF Table A89). Correction confidence was likewise largely stable, with significant changes after Holm correction occurring only in the disapproval condition in Austria and Poland (OSF Table A90). No significant round differences remained after Holm correction for perceived misinformation accuracy or fact-checking intentions (OSF Tables A91–A92).

Second, we re-estimated the main path model using mean scores across both rounds for all round-specific variables rather than relying exclusively on the Round 2 assessments (OSF Tables A93–A94). The central pattern of findings was reproduced. Relative to angry feedback, disapproving feedback predicted lower emotional exhaustion in all four countries, while approving and enthusiastic feedback predicted lower exhaustion in Austria, Denmark, and Poland, but not Bulgaria. Emotional exhaustion continued to positively predict misinformation correction avoidance in all four countries, and the positive interaction between emotional exhaustion and correction confidence remained significant in Austria, Denmark, and Poland but not Bulgaria. The pattern of indirect effects was also largely robust (OSF Table A94). Overall, these analyses indicate that the central substantive findings do not depend on the exclusive use of Round 2 measures.

Additionally, analyses using disapproval and enthusiasm as alternative reference groups confirmed that disapproval, approval, and enthusiasm did not significantly differ from one another in predicting emotional exhaustion or in their indirect effects on correction avoidance (OSF Tables A95–A98). Significant direct and indirect effects emerged exclusively for contrasts involving anger, replicating the main analysis.

5. Discussion

This study examined how the emotional tone of others' reactions to misinformation correction shapes correctors' emotional responses and subsequent behavior, an issue with direct implications for the sustainability of citizen-driven fact-checking online. Drawing on a cross-national experiment with 2001 participants from Austria, Bulgaria, Denmark, and Poland, we manipulated feedback tone across four conditions and traced its effects on emotional exhaustion and, in turn, future correction avoidance. Across countries, the findings reveal a clear and largely consistent pattern. Angry feedback uniquely depleted correctors' emotional resources, whereas other forms of feedback, including negative but less emotionally intense reactions, did not produce comparable levels of exhaustion. This emotional exhaustion, in turn, consistently predicted reduced willingness to engage in future misinformation correction. Notably, although individuals with higher correction confidence were generally less likely to disengage, confidence did not buffer

against exhaustion as expected. Instead, it amplified the relationship between exhaustion and avoidance. Taken together, these findings highlight emotional exhaustion as a central psychological cost of corrective engagement and indicate that anger-laden feedback plays a key role in this process by consistently increasing correctors' exhaustion. Thus, limiting hostile, high-arousal feedback is crucial for protecting users' emotional resources and sustaining citizen-led fact-checking, thereby supporting healthy democratic discourse online.

5.1. The distinct depleting effect of angry feedback tone

Consistent with all assumptions of H1, angry feedback induced higher emotional exhaustion than disapproving, approving, and enthusiastic feedback in Austria, Denmark, and Poland, confirming that anger-laden responses are uniquely depleting relative to all other feedback types. The additional factorial analysis further clarifies why angry feedback was particularly depleting. Rather than revealing uniform main effects of either valence or arousal, the analyses showed a significant interaction between the two dimensions across all four countries. High arousal increased exhaustion when feedback was negative, but not when feedback was positive; conversely, negative versus positive valence differentiated exhaustion under high arousal, but not under low arousal. This pattern suggests that neither negativity nor emotional activation alone sufficiently accounts for exhaustion. Instead, their combination in angry feedback appears particularly resource-depleting. This aligns with COR theory (Hobfoll, 2001) and prior work linking high-arousal negative emotions to greater resource loss (Li et al., 2017; Shah et al., 2022). Bulgaria deviated from this pattern, with only disapproving feedback differing significantly from anger; this exception is addressed below.

5.2. Exhaustion as the pathway to correction avoidance

Supporting H2, emotional exhaustion consistently predicted avoidance of misinformation correction across all four country samples. This aligns with COR theory (Hobfoll, 2001), which argues that individuals experiencing resource depletion tend to prioritize the preservation of remaining resources. Further, these results extend existing work linking emotional exhaustion to withdrawal behavior on social media (Maier et al., 2012; Sheng et al., 2023; Zhao, 2024) by identifying misinformation correction as a specific behavioral domain in which such exhaustion-driven disengagement operates. The cross-national consistency of this effect, observed across all four countries, further supports its generalizability.

Emotional exhaustion selectively mediated the relationship between feedback tone and misinformation correction avoidance, partially supporting H3. The indirect effect was robust only for disapproval, replicating consistently across all four countries. Participants who received disapproving rather than angry feedback experienced less emotional exhaustion, which in turn reduced their tendency to avoid future correction. This is consistent with COR theory (Hobfoll, 2001), which predicts that resource conservation motives triggered by emotional depletion lead to behavioral withdrawal. That the indirect effect did not extend consistently to approval or enthusiasm aligns with the results of H1, where the disapproval–anger contrast in exhaustion was the most robust cross-national finding. Notably, total effects were non-significant across all conditions and countries, suggesting that no feedback tone exerts a direct influence on correction avoidance independently of the exhaustion pathway. The robustness of the indirect effect of disapproval may reflect the role of expectancy congruence (Burgoon, 2015). Social media users tend to anticipate hostility or disagreement when engaging in interactions with others online (Vraga et al., 2015). Disapproving feedback, negative but restrained in tone, confirms this expectation without escalating it, minimizing the need for cognitive reappraisal and thereby preserving emotional resources. Positive feedback, by contrast, constitutes an expectancy violation even when pleasant, and processing

such violations may itself consume emotional resources, attenuating the restorative potential of approval and enthusiasm (Burgoon, 2015).

5.3. Correction confidence: protective alone, costly under exhaustion

Correction confidence negatively predicted misinformation correction avoidance in Austria, Denmark, and Poland, supporting H4a. Because correction confidence was assessed following feedback exposure, these findings concern participants' post-feedback confidence in their correction rather than a pre-existing efficacy disposition. H4b was not supported in any country. Rather than buffering the effect of emotional exhaustion on correction avoidance, correction confidence amplified the behavioral implications of exhaustion in Austria, Denmark, and Poland. In other words, as exhaustion rose, the protective effect of confidence was progressively undermined. We interpret this through two complementary processes described by COR theory (Hobfoll, 2001; Hobfoll et al., 2018). We hypothesized a buffering effect (H4b) because efficacy beliefs constitute a personal resource, and such resources can protect against the consequences of resource depletion. Confidence did reduce avoidance overall (H4a), but this protection did not extend to buffering the effect of exhaustion on avoidance. Under exhaustion, the pattern reversed against our prediction, and the pattern may instead reflect a second process described by COR. Confidence may not only function as a personal resource but may also be associated with greater perceived investment in the correction, given that efficacy beliefs are linked to increased effort, persistence, and goal commitment (Bandura, 1997; Honicke & Broadbent, 2016). When resources are depleted, greater perceived investment may increase the perceived costs of continued engagement, as there is more at stake to lose. Given that resource loss is particularly salient within COR, highly confident correctors may therefore become more inclined to withhold further investment and conserve their remaining resources as exhaustion increases (Hobfoll, 2001; Hobfoll et al., 2018). This interpretation is consistent with evidence that higher self-efficacy can intensify rather than protect against social media fatigue (Bright et al., 2015). Bulgaria again constituted an exception, with no significant results for either H4a or H4b.

5.4. Cross-national stability and Bulgaria's deviation

Although the paths from approving and enthusiastic feedback (vs. anger) to exhaustion were non-significant in the Bulgarian sample, unlike in Austria, Denmark, and Poland, multi-group comparisons (OSF Tables A81-A84) showed that none of Bulgaria's feedback-to-exhaustion paths differed significantly from those of the other countries, supporting the cross-national generalizability of this central mechanism (Kline, 2023).

The one path that differed significantly across countries within multi-group comparisons was the moderating role of correction confidence, which amplified the exhaustion-to-avoidance link in Austria, Denmark, and Poland but had no effect in Bulgaria. We conceptualized correction confidence as a form of self-efficacy (Bandura, 1997). The behavioral implications of efficacy beliefs vary across cultures. Intrapersonal agency beliefs, including self-efficacy, predict behavioral intentions more strongly in individualistic than in collectivistic contexts (Cho & Lee, 2015), consistent with the view that in more collectivistic settings, self-relevant psychological experiences are construed as relational rather than as intrapersonal states (Mesquita, 2001). Therefore, where behavior is guided more by relational considerations than by private efficacy, efficacy beliefs may exert less influence on how individuals respond to resource depletion.

However, this cultural account cannot stand alone, since Bulgaria and Poland both range at the collectivistic end of our sample (Minkov & Kaasa, 2022), but the effect was only absent in Bulgaria. The two diverge on structural dimensions. Bulgaria shows lower media trust and disinformation resilience, and more fully state-controlled public media,

whereas Poland is comparatively more resilient, with public-service media transitioning away from prior government control (Bleyer-Simon et al., 2025). In an environment characterized by low institutional trust and repeated failures of institutional responses to misinformation (Bleyer-Simon et al., 2025), individual corrective efforts may be perceived as less consequential, further dampening the moderating role of confidence.

Since the present design did not measure these cultural and structural factors, this interpretation remains tentative, and future research should identify what accounts for the role of correction confidence under exhaustion across differing cultural and institutional contexts.

5.5. Theoretical contributions

This study makes three theoretical contributions. First, it extends COR theory (Hobfoll, 2001) to the context of online misinformation correction. Prior applications of COR theory in digital media research have focused primarily on general social media use (e.g., Sheng et al., 2023), whereas the present findings demonstrate that the resource-loss logic applies to the specific and consequential communicative act of correcting misinformation. Angry feedback consistently depleted emotional resources, and emotional exhaustion, in turn, robustly predicted withdrawal from future correction, identifying emotional exhaustion as a critical mechanism in understanding why individuals disengage from corrective behavior online. Second, our study complicates Social Cognitive Theory's (Bandura, 1997) prediction that self-efficacy reduces the behavioral consequences of strain. The finding that correction confidence amplifies exhaustion-driven withdrawal suggests that the buffering role of efficacy beliefs may not generalize across all contexts. Future research should examine under which conditions self-efficacy protects against versus intensifies the consequences of emotional exhaustion. Third, cross-national replication of the core effects supports the generalizability of the proposed model. At the same time, the deviations in Bulgaria highlight the contextual limits and the need for culturally sensitive theorizing of conditions under which the proposed dynamics operate.

5.6. Implications

The study carries practical relevance for platform governance and intervention design. The findings suggest that the emotional climate surrounding corrective interactions plays a critical role in whether individuals remain willing to challenge misinformation. Importantly, not all critical feedback has the same psychological consequences. Angry feedback is uniquely depleting, whereas disapproval is not, indicating that tone rather than mere negativity drives emotional harm. This highlights the importance of platform moderation targeting hostile or aggressive comments in online discussions. Reducing the exposure to such reactions through platform interventions and community guidelines may help preserve the emotional resources that sustain corrective engagement while allowing critical but civil feedback. At the individual level, correction confidence reduced avoidance, suggesting that confidence-building interventions may help sustain engagement, though this effect reverses once individuals are already exhausted. Platform moderation that limits exposure to hostile feedback may therefore be a necessary precondition for such interventions to be effective. Thus, sustaining misinformation correction online depends mainly on the social conditions that shape how corrective efforts are received.

5.7. Limitations

Several limitations should be considered. First, participants were exposed to a consistent feedback tone across correction attempts, whereas real-world correctors are likely to encounter varying feedback tones over time, producing emotional dynamics that may differ from those observed in this study. Because our design cannot capture how the

observed dynamics develop through such ongoing experiences, longitudinal research is needed to examine whether exhaustion accumulates and how its effects on disengagement develop over time. Second, the feedback was presented as reactions from fellow citizens to enhance authenticity. This framing may also have influenced participants' perceptions of the feedback source, which we did not assess. Future research could examine whether responses to feedback tone vary as a function of the perceived source of this feedback. Third, our stimulus validation relied on perceived discrete emotions rather than direct ratings of valence and arousal. Although the four feedback sets were derived from the circumplex quadrants (Russell, 1980) and differed as predicted on the emotions marking those quadrants, we did not measure valence and arousal as separate dimensions and therefore cannot isolate their independent contributions to exhaustion. Future research should validate comparable stimuli directly on these dimensions. Fourth, our design manipulated the emotional tone of feedback but did not include a no-feedback control condition. Thus, although we can establish that approving and enthusiastic feedback are associated with less exhaustion than anger, we cannot determine whether they actively replenish psychological resources or simply result in less resource depletion. Future studies should include an experimental no-feedback control condition to distinguish between these possibilities. Fifth, emotional exhaustion was measured rather than experimentally manipulated, which limits causal inference about its mediating role. Sixth, correction confidence was assessed after participants had received the experimental feedback rather than before the manipulation. Although additional analyses showed no robust or consistent effect of feedback condition on correction confidence across countries, the measure should therefore be interpreted as a post-feedback appraisal rather than as a pre-existing individual characteristic unaffected by the experimental experience. Consequently, the moderation analysis establishes that the association between emotional exhaustion and correction avoidance varies as a function of post-feedback correction confidence, but it cannot determine whether the same pattern would emerge for baseline confidence assessed prior to feedback. Future research should measure correction confidence both before and after feedback to distinguish pre-existing confidence from feedback-related changes. Seventh, correction avoidance was measured as a behavioral intention rather than actual behavior. Although intentions are predictive of subsequent action (Sheeran, 2002), whether the emotional dynamics identified translate into actual corrective behavior remains to be tested through behavioral indicators to observe patterns of actual behavioral withdrawal. Eighth, all experimental materials were based on a single misinformation topic. Whether effects generalize across topics varying in political salience, topic familiarity, or polarization remains to be tested. Ninth, participants were experimentally placed into a correction scenario and told in advance that the posts contained misinformation, whereas real-world users self-select when they correct and may thus differ in confidence, motivation, or conflict tolerance. Future research should examine whether these processes are interconnected by allowing participants to assess content accuracy and decide independently whether to intervene, and then examining whether and how these decisions shape responses to feedback. Finally, while the study spans four European countries, cultural generalizability beyond this set remains uncertain.

6. Conclusion

This study shifts the focus of misinformation research from the effectiveness of corrections toward the emotional well-being of the corrector as a critical factor in sustaining online discourse. Across four countries, the findings show that emotional costs depend strongly on the tone of feedback. Specifically, angry, hostile reactions emerged as uniquely depleting, whereas both positive feedback and negative but calmly expressed disapproval did not produce comparable levels of exhaustion. The resulting emotional exhaustion serves as a primary psychological barrier that consistently leads users to avoid future

corrective engagement in order to preserve their remaining resources. These findings extend COR theory to the specific act of misinformation correction, establishing emotional exhaustion as both an outcome of hostile interactions and a driver of corrective withdrawal. Notably, while correction confidence generally reduces avoidance, it does not buffer against the effects of exhaustion and can even amplify withdrawal among those who are already depleted. This challenges traditional assumptions that self-efficacy uniformly mitigates strain and suggests that the most capable correctors may be precisely those most at risk of disengagement under hostile conditions. These results carry important practical implications. Individual motivation or confidence-building alone is insufficient if the social environment remains threatening. Instead, platform-level moderation that limits hostile and aggressive interactions is essential to protect users' emotional well-being. Ultimately, the sustainability of citizen-led fact-checking, and by extension the quality of political discourse online, depends on cultivating a more civil emotional climate on social media, as the resilience of democratic debate is closely tied to the emotional capacity of the individuals who defend it.

CRediT authorship contribution statement

Simone Moser: Writing – review & editing, Writing – original draft, Visualization, Methodology, Formal analysis, Conceptualization. **Stephanie Bühner:** Writing – review & editing, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Diana Kostikova:** Writing – review & editing, Project administration, Methodology. **Jörg Matthes:** Writing – review & editing, Supervision, Resources, Methodology, Funding acquisition, Conceptualization.

Ethics declaration

Written informed consent to take part in the study and to publish the article has been obtained from all participants or their legal representatives. The privacy rights of participants have been observed.

This study was performed in compliance with relevant laws, regulatory frameworks and guidelines where the research took place. This study was approved by the Institutional Review Board of the Faculty of Social Sciences at the University of Vienna. (Approval No. 20251110_082)

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the first author used Claude, Grammarly, and NotebookLM to improve language and clarity and to support literature synthesis. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

ChatGPT and DeepL were used to assist in translating experimental materials into German, Bulgarian, Danish, and Polish. Translations were reviewed and verified by native speakers.

Funding statement

This research was supported by the Horizon Europe programme through the project *Unveiling Emotional Dimensions of Politics to Foster European Democracy* (ENCODE; Grant Agreement ID: 101132698).

Declaration of competing interests

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chb.2026.109179>.

Data availability

Data, R scripts and outputs, as well as additional tables and figures, are openly available at <https://osf.io/5hdqg>.

References

- Aimeur, E., Amri, S., & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: A review. *Social Network Analysis and Mining*, 13(1), Article 30. <https://doi.org/10.1007/s13278-023-01028-5>
- Bandura, A. (1997). *Self efficacy: The exercise of control*. W. H. Freeman.
- Bartsch, Anne, Neuberger, Christoph, Stark, Birgit, Karnowski, Veronika, Maurer, Marcus, Pentzold, Christian, Quandt, Thorsten, Quiring, Oliver, & Schemer, Christian (2024). Epistemic authority in the digital public sphere. An integrative conceptual framework and research agenda. *Communication Theory*, 35(1), 37–50. <https://doi.org/10.1093/ct/ctae020>. <https://academic.oup.com/ct/article/35/1/37/7876430>. (Accessed 5 November 2024)
- Bautista, J. R., Zhang, Y., & Gwizdzka, J. (2022). Predicting healthcare professionals' intention to correct health misinformation on social media. *Telematics and Informatics*, 73, Article 101864. <https://doi.org/10.1016/j.tele.2022.101864>
- Bene, M., Ceron, A., Fenoll, V., Haßler, J., Kruschinski, S., Larsson, A. O., Magin, M., Schlosser, K., & Wurst, A.-K. (2022). Keep them engaged! investigating the effects of self-centered social media communication style on user engagement in 12 European countries. *Political Communication*, 39(4), 429–453. <https://doi.org/10.1080/10584609.2022.2042435>
- Bleyer-Simon, K., Horowitz, M. A., Botan, M., Brautovic, M., Bucholtz, I., Culloty, E., Gavurova, B., Grabovac, A., Kermer, J. E., Kiely, K., Leonidou, P., Mallia, M., Nenadić, I., Paisana, M., Palmer, M., Reviglio, U., Magallon, R., Salamanos, N., Aliaga, R. S., ... Weikmann, T. (2025). *How is disinformation addressed in the member states of the European Union? – 27 country cases*. European Digital Media Observatory. <https://edmo.eu/wp-content/uploads/2025/06/EDMO-Report-How-is-disinformation-addressed-in-the-member-states-of-the-European-Union-27-country-cases.pdf>
- Bode, L. (2020). User correction as a tool in the battle against social media misinformation. *Georgetown Law Technology Review*, 4(3), 367–378. <https://georgetownlawtechreview.org/wp-content/uploads/2020/07/4.2-p367-378-Bode.pdf>
- Bode, L., & Vraga, E. K. (2018). See something, say something: Correction of global health misinformation on social media. *Health Communication*, 33(9), 1131–1140. <https://doi.org/10.1080/10410236.2017.1331312>
- Bode, L., & Vraga, E. K. (2021). People-powered correction: Fixing misinformation on social media. In H. Tumber, & S. Waisbord (Eds.), *The routledge companion to media disinformation and populism* (1st ed., pp. 498–506). Routledge.
- Bright, L. F., Kleiser, S. B., & Grau, S. L. (2015). Too much Facebook? An exploratory examination of social media fatigue. *Computers in Human Behavior*, 44, 148–155. <https://doi.org/10.1016/j.chb.2014.11.048>
- Burgoon, J. K. (2015). Expectancy violations theory. In C. R. Berger, & S. R. Wilson (Eds.), *The international encyclopedia of interpersonal communication* (pp. 1–9). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118540190.wbeic102>
- Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, 28(11), 1531–1546. <https://doi.org/10.1177/0956797617714579>
- Chen, G. M., & Ng, Y. M. M. (2017). Nasty online comments anger you more than me, but nice ones make me as happy as you. *Computers in Human Behavior*, 71, 181–188. <https://doi.org/10.1016/j.chb.2017.02.010>
- Cho, H., & Lee, J.-S. (2015). The influence of self-efficacy, subjective norms, and risk perception on behavioral intentions related to the H1N1 flu pandemic: A comparison between Korea and the US. *Asian Journal of Social Psychology*, 18(4), 311–324. <https://doi.org/10.1111/ajsp.12104>
- Coppedge, M., Gerring, J., Knutsen, C. H., Lindberg, S. I., Teorell, J., Altman, D., Angiolillo, F., Bernhard, M., Cornell, A., Gjerløw, H., Glynn, A., Grahn, S., Hicken, A., Kinzelbach, K., Marquardt, K., McMann, K., Mechkova, V., Medzhorsky, J., Neundorff, A., ... Sundström, A. (2025). V-Dem dataset v15. *Varieties of Democracy (V-Dem) Project*. <https://doi.org/10.23696/VDEMDS25>
- Dailey, R. M. (2023). Confirmation in personal relationships. *Current Opinion in Psychology*, 52, Article 101593. <https://doi.org/10.1016/j.copsyc.2023.101593>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, 40(1), 55–75. <https://doi.org/10.1146/annurev-soc-071913-043137>
- Ecker, U. K., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13–29. <https://doi.org/10.1038/s44159-021-00006-y>
- Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures: Inter- and intrainational differences. *Journal of Personality and Social Psychology*, 81(5), 869–885. <https://doi.org/10.1037/0022-3514.81.5.869>
- Emerson, D. J., Hair, J. F., & Smith, K. J. (2023). Psychological distress, burnout, and business student turnover: The role of resilience as a coping mechanism. *Research in Higher Education*, 64(2), 228–259. <https://doi.org/10.1007/s11162-022-09704-9>
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. *Journalism & Mass Communication Quarterly*, 77(3), 515–540. <https://doi.org/10.1177/107769900007700304>
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218–226. <https://doi.org/10.1037/0003-066X.56.3.218>
- Greszki, R., Meyer, M., & Schoen, H. (2014). The impact of speeding on data quality in nonprobability and freshly recruited probability-based online panels. In M. Callegaro, R. Baker, J. Bethlehem, A. S. Göritz, J. A. Krosnick, & P. J. Lavrakas (Eds.), *Online panel research* (pp. 238–262). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118763520.ch11>
- Gurgun, S., Arden-Close, E., Phalp, K., & Ali, R. (2025). User correction of misinformation on social media: Perceived and actual social norms. In J. Grabis, T. E. J. Vos, M. J. Escalona, & O. Pastor (Eds.), *Research challenges in information science* (pp. 331–346). Nature Switzerland: Springer. https://doi.org/10.1007/978-3-031-92474-3_20
- Han, B. (2018). Social media burnout: Definition, measurement instrument, and why we care. *Journal of Computer Information Systems*, 58(2), 122–130. <https://doi.org/10.1080/08874417.2016.1208064>
- Hayes, A. F., Matthes, J., & Eveland, W. P. (2013). Stimulating the quasi-statistical organ: Fear of social isolation motivates the quest for knowledge of the opinion climate. *Communication Research*, 40(4), 439–462. <https://doi.org/10.1177/0093650211428608>
- Hobfoll, S. E. (2001). The influence of culture, community, and the nested-self in the stress process: Advancing conservation of resources theory. *Applied Psychology*, 50(3), 337–421. <https://doi.org/10.1111/1464-0597.00062>
- Hobfoll, S. E., Halbesleben, J., Neveu, J.-P., & Westman, M. (2018). Conservation of resources in the organizational context: The reality of resources and their consequences. *Annual Review of Organizational Psychology and Organizational Behavior*, 5, 103–128. <https://doi.org/10.1146/annurev-orgpsych-032117-104640>
- Hong, J.-C., Zhang, H.-L., Ye, J.-H., & Ye, J.-N. (2021). The effects of academic self-efficacy on vocational students behavioral engagement at school and at firm internships: A model of engagement-value of achievement motivation. *Education Sciences*, 11(8), Article 387. <https://doi.org/10.3390/educsci11080387>
- Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review*, 17, 63–84. <https://doi.org/10.1016/j.edurev.2015.11.002>
- Jerit, J., & Zhao, Y. (2020). Political misinformation. *Annual Review of Political Science*, 23, 77–94. <https://doi.org/10.1146/annurev-polisci-050718-032814>
- Johnson, G., & Connelly, S. (2014). Negative emotions in informal feedback: The benefits of disappointment and drawbacks of anger. *Human Relations*, 67(10), 1265–1290. <https://doi.org/10.1177/0018726714532856>
- Kline, R. B. (2023). Comparing groups. In *Principles and practice of structural equation modeling* (5th ed., pp. 204–213). The Guilford Press.
- Koc, M., & Barut, E. (2016). Development and validation of new media literacy scale (NMLS) for university students. *Computers in Human Behavior*, 63, 834–843. <https://doi.org/10.1016/j.chb.2016.06.035>
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>
- Lee, T., & Shi, D. (2021). A comparison of full information maximum likelihood and multiple imputation in structural equation modeling with missing data. *Psychological Methods*, 26(4), 466–485. <https://doi.org/10.1037/met0000381>
- Li, F., Wang, G., Li, Y., & Zhou, R. (2017). Job demands and driving anger: The roles of emotional exhaustion and work engagement. *Accident Analysis & Prevention*, 98, 198–205. <https://doi.org/10.1016/j.aap.2016.10.013>
- Maier, C., Laumer, S., Eckhardt, A., & Weitzel, T. (2012). When social networking turns to social overload: Explaining the stress, emotional exhaustion, and quitting behavior from social network sites' users. *ECIS 2012 Proceedings*, 71. <https://aisel.aisnet.org/ecis2012/71>
- Margova, R., & Dobrev, M. (2023). *Disinformation landscape in Bulgaria*. EU DisinfoLab. https://www.disinfo.eu/wp-content/uploads/2023/06/20230627_BulgariaDisinfoFS.pdf
- Martel, C., Pennycook, G., & Rand, D. G. (2020). Reliance on emotion promotes belief in fake news. *Cognitive Research: Principles and Implications*, 5(1), 47. <https://doi.org/10.1186/s41235-020-00252-3>
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Organizational Behavior*, 2(2), 99–113. <https://doi.org/10.1002/job.4030020205>
- Maslach, C., Jackson, S. E., & Leiter, M. P. (1997). Maslach burnout inventory: Third edition. In C. P. Zalaquett, & R. J. Wood (Eds.), *Evaluating stress: A book of resources* (pp. 191–218). Scarecrow Education.
- Matsumoto, D., Yoo, S. H., & Fontaine, J. (2008). Mapping expressive differences around the world: The relationship between emotional display rules and individualism versus collectivism. *Journal of Cross-Cultural Psychology*, 39(1), 55–74. <https://doi.org/10.1177/0022022107311854>
- Mesquita, B. (2001). Emotions in collectivist and individualist contexts. *Journal of Personality and Social Psychology*, 80(1), 68–74. <https://doi.org/10.1037/0022-3514.80.1.68>
- Minkov, M., & Kaasa, A. (2022). Do dimensions of culture exist objectively? A validation of the revised minkov-hofstede model of culture with world values Survey items and

- scores for 102 countries. *Journal of International Management*, 28(4), Article 100971. <https://doi.org/10.1016/j.intman.2022.100971>
- Muhammed, T. S., & Mathew, S. K. (2022). The disaster of misinformation: A review of research in social media. *International Journal of Data Science and Analytics*, 13(4), 271–285. <https://doi.org/10.1007/s41060-022-00311-6>
- Olley, W. O., & Ikerodah, J. O. (2025). Emotional and social barriers to engaging in misinformation correction on social media: A qualitative study. *American Journal of Arts and Human Science*, 4(3), 142–148. <https://doi.org/10.54536/ajahs.v4i3.5683>
- Prati, G., Pietrantoni, L., & Cicognani, E. (2010). Self-efficacy moderates the relationship between stress appraisal and quality of life among rescue workers. *Anxiety, Stress & Coping*, 23(4), 463–470. <https://doi.org/10.1080/10615800903431699>
- Reporters Without Borders. (2025). World press freedom index. <https://rsf.org/en/index>
- Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
- Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5), 805–819. <https://doi.org/10.1037/0022-3514.76.5.805>
- Saidah, L. N., & Hariyadi, S. (2023). From doubt to digital presence: Insecurity and self-confidence on intentions to use social media in students. *AL-ISHLAH: Jurnal Pendidikan*, 15(4), 6525–6533. <https://doi.org/10.35445/alishlah.v15i4.4794>
- Schäfer, C. (2023). *Disinformation landscape in Austria*. EU DisinfoLab. https://www.disinfo.eu/wp-content/uploads/2023/06/20230621_AustriaDisinfoFS.pdf
- Shah, S. H. A., Haider, A., Jindong, J., Mumtaz, A., & Rafiq, N. (2022). The impact of job stress and state anger on turnover intention among nurses during COVID-19: The mediating role of emotional exhaustion. *Frontiers in Psychology*, 12, Article 810378. <https://doi.org/10.3389/fpsyg.2021.810378>
- Sheeran, P. (2002). Intention—Behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1), 1–36. <https://doi.org/10.1080/14792772143000003>
- Sheng, N., Yang, C., Han, L., & Jou, M. (2023). Too much overload and concerns: Antecedents of social media fatigue and the mediating role of emotional exhaustion. *Computers in Human Behavior*, 139, Article 107500. <https://doi.org/10.1016/j.chb.2022.107500>
- Sosnowska, J., De Fruyt, F., & Hofmans, J. (2019). Relating neuroticism to emotional exhaustion: A dynamic approach to personality. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02264>
- Stsiampkouskaya, K., Joinson, A., Piwek, L., & Ahlbom, C.-P. (2021). Emotional responses to likes and comments regulate posting frequency and content change behaviour on social media: An experimental study and mediation model. *Computers in Human Behavior*, 124, Article 106940. <https://doi.org/10.1016/j.chb.2021.106940>
- Szczygiel, D., & Bazińska, R. (2013). Emotional intelligence as a moderator in the relationship between negative emotions and emotional exhaustion among employees in service sector occupations. *Polish Psychological Bulletin*, 44(2), 201–212. <https://doi.org/10.2478/ppb-2013-0023>
- The Culture Factor Group. (2023). Country comparison tool. <https://www.theculturefactor.com/country-comparison-tool>
- Tolz, Vera, Hutchings, Stephen, Kazakov, Vitaly, & Tipaldou, Sofia (2025). Disinformation as process: modeling the lifecycle of deceit. *Communication Theory*, 35(4), 214–225. <https://doi.org/10.1093/ct/ctaf009>. <https://academic.oup.com/ct/article/35/4/214/8151020>. (Accessed 26 May 2025)
- Vishkin, A., Kitayama, S., Berg, M. K., Diener, E., Gross-Manos, D., Ben-Arieh, A., & Tamir, M. (2023). Adherence to emotion norms is greater in individualist cultures than in collectivist cultures. *Journal of Personality and Social Psychology*, 124(6), 1256–1276. <https://doi.org/10.1037/pspi0000409>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Vraga, E. K., Thorson, K., Kligler-Vilenchik, N., & Gee, E. (2015). How individual sensitivities to disagreement shape youth political expression on Facebook. *Computers in Human Behavior*, 45, 281–289. <https://doi.org/10.1016/j.chb.2014.12.025>
- Weismueller, J., Gruner, R. L., Harrigan, P., Coussement, K., & Wang, S. (2024). Information sharing and political polarisation on social media: The role of falsehood and partisanship. *Information Systems Journal*, 34(3), 854–893. <https://doi.org/10.1111/isj.12453>
- Woods, F. A., & Ruscher, J. B. (2021). ‘Calling-out’ vs. ‘calling-in’ prejudice: Confrontation style affects inferred motive and expected outcomes. *British Journal of Social Psychology*, 60(1), 50–73. <https://doi.org/10.1111/bjso.12405>
- Zadroga, M. (2025). *Disinformation landscape in Poland* (No. V2). EU DisinfoLab. http://www.disinfo.eu/wp-content/uploads/2025/10/20251027_Disinfo-landscape-in-Poland-V2.pdf
- Zhao, G. (2024). Emotional exhaustion weakens the relationship between social media use and knowledge sharing behavior. *Acta Psychologica*, 250, Article 104496. <https://doi.org/10.1016/j.actpsy.2024.104496>
- Zheng, H., & Ling, R. (2021). Drivers of social media fatigue: A systematic review. *Teleatics and Informatics*, 64, Article 101696. <https://doi.org/10.1016/j.tele.2021.101696>
- Zhong, J., Shao, L., Xin, Z., & Zhang, J.-E. (2023). The mediating role of self-efficacy in the relationship between social support and work withdrawal behavior: A cross-sectional study among young lung cancer survivors. *Asia-Pacific Journal of Oncology Nursing*, 10(4), Article 100207. <https://doi.org/10.1016/j.apjon.2023.100207>

Simone Moser is a PhD student at the Department of Communication, University of Vienna, Austria. Email: s.moser@univie.ac.at. Phone: +43-1-4277-49312. ORCID: <https://orcid.org/0009-0000-3638-0015>

Stephanie Bühner is a PhD student at the Department of Communication, University of Vienna, Austria. Email: stephanie.buehrer@univie.ac.at. Phone: +43-1-4277-49342. ORCID: <https://orcid.org/0000-0003-4678-8975>

Diana Kostikova is a PhD student at the Department of Communication, University of Vienna, Austria. Email: diana.kostikova@univie.ac.at. Phone: +43-1-4277-49318. ORCID: <https://orcid.org/0009-0007-4271-1766>

Jörg Matthes is Professor of Communication at the Department of Communication, University of Vienna, Austria. Email: joerg.matthes@univie.ac.at. Phone: +43-1-4277-49307. ORCID: <https://orcid.org/0000-0001-9408-955X>