

I Reason Who I am? Identity Salience Manipulation to Reduce Motivated Reasoning in News Consumption

Magdalena Wischniewski *
Department of Social Psychology, University of
Duisburg-Essen, Germany
magdalena.wischniewski@uni-due.de

Nicole Krämer
Department of Social Psychology, University of
Duisburg-Essen, Germany
nicole.kraemer@uni-due.de

ABSTRACT

Past research has drawn on motivated reasoning theories in order to explain why some people fall for fake news while others do not. One such motivated reasoning paradigm proposes an elicitation of identity threat when incoming information is inconsistent with prior attitudes and beliefs. This experienced identity threat leads to biased information processing in order to defend those prior attitudes and beliefs. Building on this, we conducted two studies to test the overarching hypothesis that shifting identity salience changes information processing outcomes. In two experimental studies with $N = 353$, we tried to (1) increase factual information acceptance and (2) decrease misinformation acceptance. Our data support the previously found results that identity-threatening information decreases the evaluation of information compared to a control group. Findings also suggested that identity-supporting information was evaluated better, respectively. However, in both studies, identity salience manipulation did not change the evaluation of the information. Still, we found that those participants for whom another identity was made more salient indicated reduced feelings of anger compared to participants who were threatened and received no identity salience manipulation. We interpret these results as a promising first step to counter motivated reasoning processes.

CCS CONCEPTS

• **Social and Professional Topics**; • **User Characteristics**; • **Cultural Characteristics**; • **Collaborative and social computing**; • **Collaborative and social computing theory, concepts and paradigms**;

KEYWORDS

Motivated Reasoning, Identity Protection Cognition, Identity Salience, Misinformation, Anger

ACM Reference Format:

Magdalena Wischniewski and Nicole Krämer. 2020. I Reason Who I am? Identity Salience Manipulation to Reduce Motivated Reasoning in News

*corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

SM Society '20, July 22–24, 2020, Toronto, ON, Canada

© 2020 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-7688-4/20/07...\$15.00

<https://doi.org/10.1145/3400806.3400824>

Consumption. In *International Conference on Social Media and Society (SM-Society '20)*, July 22–24, 2020, Toronto, ON, Canada. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3400806.3400824>

1 INTRODUCTION

Today, misinformation spreads online with rapid velocity and seems to become ever more pervasive [42]. Repercussions of this vast increase of shared misinformation have not only become tangible in the health sector (see outbreaks of measles) [30] but more generally relate to increased distrust of public institutions [28]. Notably, social media sites such as Facebook, Twitter, and Reddit, that facilitate the creation and sharing of user-generated content are often associated with dissemination and consumption of misinformation [29]. In addition to that, messaging services like WhatsApp contribute increasingly to the circulation of misinformation [31]. These developments coincide with two facilitating processes, namely, (1) increasing numbers of people receive their news through social media platforms [32] and (2) content is created and spread by automated accounts (e.g., social bots) [43] and so-called trolls (e.g., Russian Intelligence Research Agency) [3] who target individual users. In all likelihood, more social media users will be exposed to misinformation content in the future.

In light of these trends, the question of why people fall for fake news has not only been raised within the research community [26] but also on popular news sites [27]. We can find answers in prior research on motivated reasoning. Motivated reasoning theories propose that individuals sometimes process incoming information with a bias in favor of existing beliefs and attitudes, “to arrive at a desired conclusion” [[2], p. 485]. According to this, motivated reasoning affects the perception of both misinformation and accurate information: misinformation that supports existing beliefs and attitudes is likely to be accepted (false positive acceptance), just as misinformation that contradicts those beliefs and attitudes is easily detected (true positive). In return, information rejection of factual information contradicting an individual’s prior beliefs and attitudes are higher (false negative) and vice-a-versa, in cases where information acceptance coincides with the individual’s prior belief’s and attitudes (true negative).

Analogous to the discussion of why people fall for fake news, studies examining the possibilities to protect individuals from misinformation, as well as to correct misinformation at its source, have increased in number. Many of these investigate misinformation corrections [8] and the acceptance of these corrections [25], but also inoculation strategies [6] and media literacy interventions [34]. This paper contributes to the existing literature by introducing another approach to deal with the increase of misinformation circulation and propose ways of protecting users. We employ a social

identity approach to address this. We propose manipulating identity salience of users' to either a) make a threatened identity less salient or b) to reduce the salience of a misinformation affirming identity.

2 THEORETICAL BACKGROUND

2.1 Information processing and motivated reasoning

The central claim of motivated reasoning theories is that the individual assessment of information is driven by prior attitudes [36] and beliefs as well as group belonging [5] and identity [16]. Motivated reasoning theories claim that if incoming information is congruent with an individual's prior beliefs and attitudes or if it supports one's world view, it is more likely to be believed, whereas information contradicting individuals' views is more likely to be rejected.

To explain the underlying psychological mechanisms of motivated reasoning, different theoretical approaches have been proposed. With origins in cognitive dissonance theory [7], earlier theories suggest that incoming information elicits different motives like the need for accuracy, self-defense and impression management [4] which are then associated with biased memory search and selection of reasoning strategies [20]. In particular, defense motivation as a result of perceived social identity threat has been related to directionally motivated information processing [13]. More recently, emotional processes have been introduced to add to the existing literature on cognitive and motivational approaches. For example, in coining the term *affective contagion*, Lodge and Taber [22], acknowledge these affective processes within motivated reasoning. They argue that early affective associations towards the incoming information determine memory retrieval and set the direction for subsequent processing via associative pathways [37]. By contrast, instead of focusing on early affective reactions, Suhay and Erisen [35] investigate the discrete emotions of anger, anxiety and enthusiasm, which all occur later, and their effect on motivated reasoning. Through mediation models, they confirmed the role of anger for attitude inconsistent information and enthusiasm for attitude consistent information. In their analysis, however, motivated reasoning was mostly fueled by anger. In addition to this research, a third theoretical approach centers on the role of "Identity Protection Cognition" (IPC) within the psychological processes of motivated reasoning.

2.2 Identity Protection Cognition and Identity Salience

A prerequisite for Identity Protection Cognition posits that specific ideas, ideologies or world views become "a badge of membership with identity-defining affinity groups" [[9], p. 2]. The formation of such groups is evident, especially when issues become politicized. In other words, group identities and alliances become tightly bound to specific belief systems which in turn affect behavior. According to Identity Protection Cognition, it is considered *individually rational* for individuals to reject information that is contrary to group beliefs, because acceptance would threaten their status within their affinity group. "[I]n fact, identity-protective cognition is a mode of engaging information rationally suited to the ends of the agents

who display it" [[9], p.1]. It is then the primary goal of an individual to protect her or his status within the respective group. In this, Identity Protection Cognition draws on both evolutionary psychology and utility maximation theory. While the former proposes that social groups fulfil basic needs of belonging, protection and safety, the latter supplements utility maximation theory in proposing that the benefits of conforming to group beliefs outweigh the costs of rejecting them. Furthermore, empirical studies in the context of politics [2] and technology acceptance [11] support Identity Protection Cognition. Van Bavel and Pereira [2] suggest in their identity-based model of beliefs that "accuracy goals compete with [partisan] identity goals to determine the value of beliefs" (p. 215). For our part, instead of focusing on partisan related identities and goals, we want to focus on the underlying processes of identity threat and identity affirmation.

As stated above, identity-based beliefs may either support or reject factual information or misinformation. In our two studies we focussed on false negatives and false positives (see *c* and *d* in Table 1). The overall aims are to increase acceptance and decrease rejection of factual information (thereby addressing false negatives), and decrease acceptance and increase rejection of misinformation (with an eye to false positives), through a manipulation of identity salience. We hypothesize that changing the salience of an individual's identity from a threatened identity to an unthreatened identity will, in turn, increase the likelihood of accepting factual information and rejecting misinformation.

In general, identity salience manipulations have been successfully implemented to change attitudes and behavior as observed in studies on stereotype susceptibility [33], stereotype threat [24] and performance [18] but also within the realm of policy support [41]. These studies draw mostly on social identity theory and self-categorization theory. According to social identity theory, people can define themselves both in terms of who they are as an individual as well as their membership in various groups [38]. Turner, Oakes, Haslam and McGarty [40] extend this view by what they called self-categorization theory. Self-categorization theory picks up the concept of personal and social identity and describes these as "different levels of self-categorization" (p.1). It asserts an experience of self through varying identities, and that the prevalent identity shifts in response to contextual and social cues. Consequently, once another identity becomes more salient, the respective norms of the salient group guide downstream cognition and emotion [15].

Moreover, we hypothesize that a shift in identity salience would not only affect how individuals think about incoming information but also how they feel about it. Because experiences of identity threat have previously been shown to induce either feelings of anxiety or anger [16], we hypothesized that a manipulation of identity salience should also be reflected by an identifiable emotional reaction. Depending on the emotion induced by the identity-threatening information (either anger or anxiety), this emotion should decrease upon the manipulation of identity salience. Building on this, we conducted two studies, (1) which tried to increase the acceptance of factual information and (2) which tried to increase the rejection of misinformation.

Table 1: Processing (mis-)information.

	Information is true (factual information)	Information is false (misinformation)
Belief supports information	(a) True positive	(b) False positive
Belief contradicts information	(c) False negative	(d) True negative

2.3 Hypotheses Study 1

In study 1, we hypothesized the following: (H1.1) If incoming information (news) threatens the identity of individuals, the information is evaluated more poorly than compared to a group whose identity is not threatened by the information. (H1.2) This identity threat is accompanied by an emotional reaction. Therefore, identity-threatened individuals experience higher levels of anger or anxieties. (H2.1) When the salient identity of an individual is changed to an unthreatened identity, the individual will evaluate the factual information better. (H2.2) An identity salience manipulation for an unthreatened identity will result in no changes in evaluation. (H3.1) The change from a threatened to an unthreatened identity will also be reflected in a change to the emotion experienced (i.e., individuals will experience less anger or anxiety). (H3.2) The identity salience manipulation for unthreatened identities will result in no changes in experienced emotion. See Appendix A.1 for an overview of all hypotheses.

2.4 Hypotheses Study 2

For study 2, instead of increasing the acceptance of factual information, we intended to increase the rejection of misinformation (i.e., (b) false positives in Table 1). Building on IPC, we hypothesized the following. (H4.1) If incoming information (misinformation) supports the identity of individuals, the information evaluation is better than compared to a group whose identity is not supported by the information. (H4.2) This identity support is accompanied by an emotional reaction towards the identity-supporting stimulus. Therefore, individuals experience higher levels of enthusiasm. (H5.1) By changing the salient identity of individuals to a non-supporting identity, the affected individuals will have a poorer evaluation of the misinformation. (H5.2) An identity salience manipulation from one non-supporting identity to another will result in no changes in evaluation. (H6.1) The change from a supporting to a non-supporting identity will also be reflected in a change of experienced emotion. Individuals should experience less enthusiasm. (H6.2) The identity salience manipulation from non-supported to another non-supported identity will result in no changes in experienced emotion. See Appendix A.2 for an overview of all hypotheses.

3 METHODS

Both studies received ethical approval by the ethics committee of the Department of Computer Science and Applied Cognitive Science, University of Duisburg-Essen.

3.1 Study 1 - Design and Procedure

To test our hypotheses, we conducted a 2 (manipulation versus no manipulation) $\times$ 2 (women versus men) between-subject study with

247 University students (175 female) with an age range of 18 to 62 years ($M = 23.36$, $SD = 3.65$). Participants of both groups (control and experimental group) were asked to read and evaluate a news article. In addition, we asked for participants' affective reactions after reading the article. After being randomly assigned to one of the two groups, participants in the experimental group received an identity salience manipulation before reading the news article. All participants were debriefed upon completion.

In order to reduce identity-threatening potential, per our hypotheses, of the news article, we implemented an identity salience manipulation. Since all participants were university students, we intended to shift identity salience to that identity. For this, we applied a manipulation used by Shi, Pittinsky, and Ambady [33], which has previously been found to induce identity salience change. Participants of the experimental group were asked the following four questions: 1) if they were enrolled at a university, 2) what subject they were studying, 3) which semester they were in, and 4) if they were living in a student home.

3.2 Stimulus material

The article presented to all participants discussed factual information about domestic violence. Specifically, it focused on domestic violence committed by female perpetrators against their male partners, and, hence, was written in a way that threatened women's identity. The article was relatively short (approximately 350 words) and was presented in the design of a well-known newspaper.

3.3 Evaluation of the articles

To measure how the participants evaluated the article, we used the Trust in News Media (TiNM) scale by Kohring and Matthes [19]. The scale is a standardized and validated multidimensional measurement to depict trust and credibility of news media. It consists of four lower-order factors that are assessed by four items: (1) trust in the selectivity of topics, (2) trust in the selectivity of facts, (3) trust in the accuracy of depictions and (4) trust in journalistic assessment. Answers were given on a five-point Likert scale (1 = "do not agree at all", 5 = "fully agree") and items reached acceptable reliability of Cronbach's $\alpha = .86$. The complete list of items can be assessed in Table 2

3.4 Emotions

To assess emotional reactions towards the news article, we asked participants to self-report their experience of anger and anxiety. To do so, we created six items based on Affective Intelligence Theory (AIT) [23]. Anger and anxiety were measured through three items respectively: hateful, angry, outraged (Cronbach's $\alpha = .65$) and afraid, worried, anxious (Cronbach's $\alpha = .71$). Participants indicated on a five-point Likert scale (1 = "do not agree at all", 5 = "fully

Table 2: Items of the Trust in News Media Scale by Kohring and Matthes.

	Items
Selectivity of topics	The topic of domestic violence by female perpetrators received the necessary attention.
	The topic of domestic violence by female perpetrator is assigned an adequate status
	The frequency with which domestic violence by female perpetrators is covered is adequate.
Selectivity of facts	The topic is covered on the necessary regular basis.
	The essential points are included.
	The focus is on important facts.
	All important information regarding the topic of domestic violence by female perpetrators is provided.
Accuracy of depiction	Reporting includes different points of view.
	The information in a report would be verifiable if examined.
	The reported information is true.
	The report recounts the facts truthfully.
Journalistic assessment	The facts that I received regarding domestic violence by female perpetrator are correct
	Criticism is expressed in an adequate manner.
	The journalist's opinions are well-founded.
	The commentary regarding domestic violence by female perpetrators consists of well-considered conclusions.
	I feel that the journalistic assessment regarding the topic of domestic violence by female perpetrators is useful.

Table 3: Mean scores of TiNM, self-reported anger and anxiety by group.

Condition		TiSM		Anger		Anxiety	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Salience manipulated	female	40.40	9.46	6.57	2.54	7.40	2.80
	male	42.50	9.24	6.30	2.05	6.65	2.56
Salience not manipulated	female	38.62	9.83	7.46	2.87	7.53	2.71
	male	43.13	10.91	6.41	3.01	6.84	2.55

agree”) how they felt when reading the article. The use of AIT to assess emotions via self-reporting is sound and has previously been used in several studies.

3.5 Results study 1

Mean scores over all groups for the Trust in News Media scale, self-reported anger and self-reported anxiety are presented in Table 3. To test the influence of identity threat on article evaluation (H1.1) and the proposed effect of an identity salience manipulation to decrease that effect (H2.1), we conducted an analysis of variance (ANOVA) with planned contrasts, with a standard $p < .05$ criterium of significance. The groups (manipulation/no manipulation) were entered as independent factors whereas the Trust in News Media score was entered as a dependent variable.

The overall model did not reach significance, $F(3, 242) = 2.6$, $p = .053$, $\eta_p^2 = .03$. Yet, to answer H1.1, we compared the evaluation of female readers with male readers through planned contrast. H1.1 hypothesized that female readers would evaluate the article worse than male readers. This was supported in our data: male readers assigned a higher quality evaluation to the article than did female

readers ($F(1, 242) = 6.22$, $p = .01$, $\eta_p^2 = .03$). However, this was neither reflected by the reported experience of anger ($F(1, 242) = 3.1$, $p = .08$, $\eta_p^2 = .01$) nor of anxiety ($F(1, 242) = 3.4$, $p = .06$, $\eta_p^2 = .01$), as we had hypothesized in H1.2.

Concerning H2.1, which hypothesized that an identity salience manipulation would lead to a decreased identity threat, we conducted another planned contrast. However, the identity salience manipulation in one group did not change how females evaluated the information compared to females of the group without a manipulation ($F(1, 242) = 1.37$, $p = .24$). As hypothesized in H2.2, we did not find a change in the evaluation between male readers ($F(1, 242) = 0.1$, $p = .76$). Although our data did not support that the identity salience manipulation changed the evaluation of the article, we found that female readers of the salience manipulation reported lower levels of anger than females of the control group ($F(1, 242) = 4.82$, $p = .03$). These data support H3.1. This was, however, only true for reported levels of anger but not for anxiety ($F(1, 242) = 0.12$, $p = .74$). As hypothesized (H3.2), male readers of the salience manipulation group did not differ from male readers of the control

Table 4: Items of the Group Identification Scale. (R) indicates a reversed-scored item.

Subscale	Item
Affective	1. I would like to be in a different group (R).
	2. Members of this group like one another.
	3. I enjoy interacting with the members of this group.
	4. I don't like many of the people in this group (R).
Behavioral	5. In this group, members don't have to rely on one another (R).
	6. All members need to contribute to achieve the group's goals
	7. This group accomplishes things that no single member could achieve alone.
	8. In this group, members do not need to cooperate to complete group tasks (R).
Cognitive	9. I think of this group as a part of who I am.
	10. I see myself as quite different from other members of the group (R).
	11. I don't think of this group as a part of who I am (R).
	12. I see myself as quite similar to other members of the group.

group concerning anger ($F(1, 242) = 0.03, p = .87$) and anxiety ($F(1, 242) = 0.1, p = .75$).

The results from study 1 showed that Identity Protection Cognition was supported. The threatened group gave a worse evaluation of the factual news article than did the non-threatened group, as was also reflected in self-reported levels of anger. The manipulation of identity salience to close this gap was unsuccessful. After receiving the manipulation, mean evaluations did not change significantly, although we did observe an increase in mean evaluation in the hypothesized direction. Self-reported levels of anger changed significantly after the manipulation, but not self-reported levels of anxiety.

3.6 Study 2 - Design and Procedure

The basic construction of study 2 was similar to study 1. To test our hypotheses, we conducted a 2 (manipulation versus no manipulation) $\times$ 2 (vegetarian versus meat-eater) between-subject study with 106 University students (75 female) with an age range of 18 to 56 years ($M = 23.44, SD = 4.01$). Participants of both groups (control and experimental group) were asked to read and evaluate a news article. In contrast to study 1, the news article was written in the style of misinformation. In doing so, we oriented ourselves to findings by Horne and Adali [14], including elements like simple language and sensationalizing headlines. After evaluating the quality of the article, participants were asked to indicate their experienced enthusiasm. As in study 1, one group's participants received an identity salience manipulation before reading the news article in order to reduce our hypothesized identity-supporting potential of the misinformation article. Here we aimed to make the identity of *student* more salient. After the experience in study 1 of employing a manipulation consisting of only four questions, we decided to use a more elaborate manipulation in study 2. We applied the group identification scale developed by Henry, Arrow, and Carini [10]. The scale consists of 12 items which can be divided into three subscales: cognitive (social categorization), affective (interpersonal attraction) and behavioral (interdependence). For a full list of all questions see Table 4. After completion, all participants were debriefed.

3.7 Stimulus Material

The article presented to all participants discussed misinformation about the apparent positive effect of a vegetarian diet on job success. It was written in such a way as to venerate individuals who are vegetarian or vegan. The article was similar in length to study 1 (~ 330 words) and was presented in the design of a well-known newspaper.

3.8 Evaluation of the article

As in study 1, we used the Trust in News Media (TiNM) scale to assess how participants evaluated the article.

3.9 Emotions

Because we expected that participants following a vegetarian diet would feel positively about the identity-supporting misinformation, instead of assessing anger and anxiety, participants were asked to self-report feelings of enthusiasm. As with the evaluation of anger and anxiety in study 1, the measure of enthusiasm was based on *Affective Intelligence Theory* (AIT) by [23], and was assessed through three categories: proud, enthusiastic and hopeful (Cronbach's $\alpha = .86$). Participants indicated on a five-point Likert scale (1 = "do not agree at all", 5 = "fully agree") how they felt when reading the article.

3.10 Results study 2

For an overview of mean responses concerning TiNM scores and self-reported enthusiasm, see Table 5. To determine whether individuals whose identity was supported by the misinformation evaluated the article better than those whose identity was not supported (H4.1), we conducted an analysis of variance (ANOVA) with planned contrasts and controlled for age and gender, with a standard $p < .05$ criterium of significance. The overall model was not significant concerning the grouping variable ($F(3, 99) = 2.63, p = .054, \eta_p^2 = .07$). However, the planned contrast revealed that vegetarian readers evaluated the misinformation significantly better than meat-eaters ($F(1, 99) = 4.66, p = .03$), as was predicted

Table 5: Mean scores of TiNM score and self-reported enthusiasm by group.

Condition		TiSM		Enthusiasm	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Salience manipulated	vegetarian	46.62	8.22	10.00	2.51
	meat eater	45.57	9.26	6.03	2.88
Salience not manipulated	vegetarian	48.75	7.78	9.88	3.27
	meat eater	42.61	7.8	6.52	2.58

by identity-support hypothesis (H4.1). This was also reflected in higher self-reported levels of enthusiasm ($F(1, 99) = 39.07, p < .001, \eta_p^2 = .28$) (H4.2).

To test whether shifting identity salience from an identity-support to an uninvolved identity, we conducted another planned contrast. However, results revealed no significant change between vegetarian readers of the manipulation group and vegetarian readers of the control group ($F(1, 99) = 1.24, p = .27, \eta_p^2 = .01$) (H5.1). Nevertheless, inspecting the mean scores (see Table 5), we noted a change in the hypothesized direction ($M_{\text{vege_manipulated}} < M_{\text{vege_control}}$). For the non-supported identity, we found that the manipulation did not change the evaluation as was expected ($F(1, 99) = 2.14, p = .15, \eta_p^2 = .02$) (H5.2). This overall picture was confirmed when running an ANOVA with planned contrast for the dependent variable enthusiasm as well. Manipulation did not affect the responses for enthusiasm of identity-supported individuals ($F(1, 99) = 0.04, p = .85, \eta_p^2 < .001$) (H6.1) and non-supported individuals ($F(1, 99) = 0.63, p = .43, \eta_p^2 = .02$) (H6.2).

4 DISCUSSION

In this paper, we wanted to assess if an individual's evaluation of a news article changes when their identity salience is manipulated. In doing so, our overall aim was to increase the acceptance of factual information on the one hand and to increase rejections of misinformation on the other. We build on Identity Protection Cognition which argues both that information is rejected when it threatens an individual's identity and that a piece of information is accepted when it supports an individual's identity. Our hypothesis was, then, that by changing an individual's identity from a threatened to an unthreatened identity, they would be more likely to accept incoming information. Further, by our changing an individual's salient identity from a supporting identity to a non-supporting, they would be more likely to reject incoming information. We applied this to increase factual information acceptance (study 1) and increase misinformation rejection (study 2).

The results of study 1 confirmed Identity Protection Cognition insofar as individuals whose identity was threatened by the information of the news article did indeed evaluate the article to be worse than individuals whose identity was not threatened. They also self-reported significantly higher instances of experiencing anger. However, even after changing identity salience to an unthreatened identity, we did not find that individuals indicated significantly increased acceptance of the news article compared to the control group who did not receive the salience manipulation. However, we did find that, although evaluation did not change after the salience manipulation, levels of self-reported anger were significantly lower

for individuals of the manipulation condition. We did not observe this finding for levels of self-reported anxiety.

The aim of study 2 was to increase misinformation rejection for individuals whose identity was supported by the misinformation. Our results confirmed Identity Protection Cognition, as individuals whose identity was supported evaluated the misinformation better than those whose identity was not supported. We found this to be reflected by the self-reported levels of enthusiasm, as individuals whose identity was supported indicated significantly higher levels of enthusiasm. Contrary to our hypotheses, changing identity salience to a non-supporting identity did not result in lower levels of misinformation acceptance. Although mean acceptance decreased, this change was not significant. Likewise, we found no support for our hypothesis in self-reported experiences of enthusiasm. The identity salience manipulation did not change how individuals felt about the misinformation.

4.1 Limitations

Before we dive deeper into implications, we want to elaborate on the limitations of our findings. In study 1, we found evidence in our data that manipulating identity salience led to lower levels of anger. This suggests that the information was less identity-threatening. However, we did not find this reflected in the mean evaluation (TiNM scores). Although the mean evaluation of participants from the salience manipulation condition was higher as hypothesized ($M_{\text{manipulation}} = 40.40; M_{\text{no manipulation}} = 38.62$), the difference was not significant. Results from study 2 confirmed this. Similarly, while the mean evaluation decreased in the hypothesized direction, the change was not significant. We identify four possibly explanations for these results.

(1) The manipulation had no real effect across the two studies. However, we would like to call attention to the significant change in experienced anger that was recorded. Given this, we assume that at least in study 1 the manipulation worked. Future studies would need to further investigate the shift in identity salience via manipulation.

(2) Identity salience was manipulated, but the effect of such a manipulation was not lasting. Again, we cannot refuse this either. We know that the duration of framing effects depend heavily on individual characteristics like a person's political knowledge [21]. Nevertheless, little is known about the effectiveness of identity salience manipulations.

(3) The manipulation worked but its effect was too weak to be statistically significant.

(4) Similarly to the previous possible explanation, both studies may have been underpowered. To this and the previous explanation,

we see reason to believe that our results support our original hypotheses. Our data demonstrated the expected direction of change for both studies when scrutinizing the mean differences, and we saw the same pattern for the experienced emotion of anger.

4.2 Implications and Conclusion

Despite its limitations, our research offers some valuable conclusions. First, both studies support claims made by Identity Protection Cognition. Threatened individuals evaluated factual information as being of poorer quality, and experienced higher levels of anger. At the same time, information, in our case misinformation, that supports an identity is evaluated as being of better quality, while also reflecting a higher level of experienced enthusiasm on the part of the reader. This has important implications for how we approach misinformation and factual information acceptance, namely, going further than confirmation and partisan bias. While it is true that issues arise when prior attitudes are confirmed or rejected by incoming information, incorporating attitudes in a model of identity might explain why misinformation corrections [8], fact-checking [1], and other attempts to rectify false information fail. If an attitude or previously held opinion becomes closely associated with an affinity group, as Identity Protection Cognition suggests, corrections to information are likely to be ignored, counterargued, or outright dismissed. Although studies have shown that individuals are willing to adapt their views [39]—a phenomenon that has previously been related to “Cognitive Reflection” [26]—we find those cases to be generally less polarized. Studies have shown that beliefs can become more polarised after increasing the saliency of the reader’s political identity [41], while others have indicated that intergroup comparisons can maximise perception of between-group differences [12]. In light of this, we argue that whether it is delivering news or offering misinformation corrections, practitioners need to consider these effects that are so associated with the readers’ identities. This implies refraining from framing issues in a politicized light but instead focusing on what is said rather than by whom.

In closing, let us turn to the by far more obvious problem: social media. Here, admittedly, little can be done concerning the visibility of political identities. We fear that other issues like content moderation [9] and the monetization of clicks and views may be far more decisive than any identity salience manipulation we might devise. Nevertheless, we do not intend to discourage anyone. In fact, building on our results, we encourage our colleagues to investigate not only the role of identity in information processing and motivated reasoning, but also emotional processes and contextual cues.

ACKNOWLEDGMENTS

This research was funded by the German Research Foundation (DFG) under grant No. GRK 2167, Research Training Group “User-Centred Social Media”. Special thanks to Louisa Nosch, Agnes Dyszlewski, Ida Schaffeld, and Britta Tröger for their help in the conceptualization and data collection.

REFERENCES

- [1] Michael J. Aird, Ullrich K.H. Ecker, Briony Swire, Adam J. Berinsky, and Stephan Lewandowsky. 2018. Does truth matter to voters? The effects of correcting political misinformation in an Australian sample. *R. Soc. Open Sci.* 5, 12 (2018). DOI:https://doi.org/10.1098/rsos.180593
- [2] J. J. Van Bavel and A. Pereira. 2018. The partisan brain: An identity-based model of political belief. *Trends Cogn. Sci.* 22, 3 (2018), 213–224.
- [3] Brandon Boatwright, Darren L. Linville, and Patrick L. Warren. 2018. Troll factories: The IRA and state-sponsored agenda building. *Resour. Cent. Media Free. Eur.* (2018).
- [4] Shelly Chaiken, Roger Giner-Sorolla, and Serena Chen. 1996. Beyond accuracy: Defense and impression motives in heuristic and systematic information processing. In *The Psychology of action: Linking cognition and motivation to behavior*, P. M. Gollwitzer and J. A. Bargh (eds.). Guilford Press, New York, NY, 553–578.
- [5] G. L. Cohen. 2003. Party over policy: The dominating impact of group influence on political beliefs. *J. Pers. Soc. Psychol.* 85, 5 (2003), 808–822. DOI:https://doi.org/10.1037/0022-3514.85.5.808
- [6] John Cook, Stephan Lewandowsky, and Ullrich K.H. Ecker. 2017. Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PLoS One* 12, 5 (2017), 1–21. DOI:https://doi.org/10.1371/journal.pone.0175799
- [7] L. Festinger. 1976. *A theory of cognitive dissonance*. Stanford University Press, Stanford, CA.
- [8] D. J. Flynn, Brendan Nyhan, and Jason Reifler. 2017. The Nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Polit. Psychol.* 38, 1 (2017), 127–150. DOI:https://doi.org/10.1111/pops.12394
- [9] Giovanni De Gregorio. 2019. Democratising online content moderation: A constitutional framework. *Comput. Law Secur. Rev.* xxx (2019), 105374. DOI:https://doi.org/10.1016/j.clsr.2019.105374
- [10] Kelly Bouas Henry, Holly Arrow, and Barbara Carini. 1999. A tripartite model of group identification: Theory and measurement. *Small Gr. Res.* 30, 5 (1999), 558–581. DOI:https://doi.org/10.1177/104649649903000504
- [11] Shirley S. Ho, Dietram A. Scheufele, and Elizabeth A. Corley. 2010. Making sense of policy choices: Understanding the roles of value predispositions, mass media, and cognitive processing in public attitudes toward nanotechnology. *J. Nanoparticle Res.* 12, 8 (2010), 2703–2715. DOI:https://doi.org/10.1007/s11051-010-0038-8
- [12] Michael A. Hogg, J. Turner, and Barbara Davidson. 1990. Polarized norms and social frames of reference: A test of the self-categorization theory of group polarization. *Basic Appl. Soc. Psychol.* 11, 1 (1990).
- [13] Natascha de Hoog. 2013. Processing of social identity threats: A defense motivation perspective. *Soc. Psychol. (Gott)* 44, (2013), 361–372. DOI:https://doi.org/10.1027/1864-9335/a000133
- [14] Benjamin D. Horne and Sibel Adali. 2017. This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. (2017), 759–766. Retrieved from <http://arxiv.org/abs/1703.09398>
- [15] Matthew J. Hornsey. 2008. Social identity theory and self-categorization theory: A historical review. *Soc. Personal. Psychol. Compass* 2, 1 (2008), 204–222. DOI:https://doi.org/10.1111/j.1751-9004.2007.00066.x
- [16] Leonie Huddy, Lilliana Mason, and Lene Aarøe. 2015. Expressive partisanship: campaign involvement, political emotion, and partisan identity. *Am. Polit. Sci. Rev.* 109, 1 (2015), 1–17. DOI:https://doi.org/10.1017/S0003055414000604
- [17] Dan M. Kahan. 2017. Misconceptions, misinformation, and the logic of identity-protective cognition.
- [18] D. Van Knippenberg and N. Ellemers. 2003. Social identity and group performance. Identification as the key to group-oriented effort. In *Social identity at work: Developing theory for organizational practice*, S. Alexander Haslam, Daan van Knippenberg, Michael J. Platow and Naomi Ellemers (eds.). Psychology Press.
- [19] Matthias Kohring and Jörg Matthes. 2007. Trust in news media: Development and validation of a multidimensional scale. *Communic. Res.* 34, 2 (2007), 231–252. DOI:https://doi.org/10.1177/0093650206298071
- [20] Ziva Kunda. 1990. The case for motivated reasoning. *Psychol. Bull.* 108, 3 (1990), 480–498. DOI:https://doi.org/10.1037/0033-2909.108.3.480
- [21] Sophie Lecheler, Andreas R.T. Schuck, and Claes H. De Vreese. 2013. Dealing with feelings: Positive and negative discrete emotions as mediators of news framing effects. *Communications* 38, 2 (2013), 189–209. DOI:https://doi.org/10.1515/commun-2013-0011
- [22] Milton Lodge and Charles S. Taber. 2013. Affective contagion and political thinking. In *The Rationalizing voter*.
- [23] George E. Marcus, W. Russell Neuman, and Michael B. MacKuen. 2000. *Affective intelligence and political judgement*. University of Chicago Press.
- [24] Matthew S. McGlone and Joshua Aronson. 2006. Stereotype threat, identity salience, and spatial reasoning. *J. Appl. Dev. Psychol.* 27, 5 (2006), 486–493. DOI:https://doi.org/10.1016/j.appdev.2006.06.003
- [25] Brendan Nyhan and Jason Reifler. 2010. When corrections fail: The persistence of political misperceptions. *Polit. Behav.* 32, 2 (2010), 303–330.
- [26] Gordon Pennycook and David G. Rand. 2018. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* (2018), 1–12. DOI:https://doi.org/10.1016/j.cognition.2018.06.011
- [27] Gordon Pennycook and David G. Rand. 2019. Why do people fall for fake news? *The New York Times*.

- [28] Nathaniel Persily. 2017. The 2016 U.S. Election: Can democracy survive the internet? *J. Democr.*28, 2 (2017), 63–76.
- [29] Pew Research Center. 2016. Many americans believe fake news is sowing confusion. *Encycl. Fam. Stud.* (2016). DOI:https://doi.org/10.1002/9781119085621.wbefs533
- [30] Jacek Radzikowski, A. Stefanidis, Kathryn H Jacobsen, Arie Croitoru, A. Crooks, and Paul L. Delamater. 2016. The measles vaccination narrative in twitter: A quantitative analysis. *JMIR Public Heal. Surveill.*2, 1 (2016), 1–15. DOI:https://doi.org/10.2196/publichealth.5059
- [31] Gustavo Resende, Philippe Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara M. Almeida, and Fabricia Benevenuto. 2019. (Mis) Information dissemination in WhatsApp: Gathering, analyzing and countermeasures. *Proc. World Wide Conf.* (2019), 818–828.
- [32] Elisa Shearer and Elizabeth Grieco. 2019. Americans are wary of the role social media sites play in delivering the news. *Pew Res. Cent.* (2019).
- [33] Margaret Shih, Todd L. Pittinsky, and Nalini Ambady. 1999. Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychol. Sci.*10, 1 (1999), 80–83. DOI:https://doi.org/10.1111/1467-9280.00111
- [34] Susan Currie Sivek. 2018. Both facts and feelings: Emotion and news literacy. *J. Media Lit. Educ.*10, 2 (2018), 123–138. DOI:https://doi.org/10.23860/jmle-2018-10-2-7
- [35] Elizabeth Suhay and Cengiz Erisen. 2018. The role of anger in the biased assimilation of political information. *Polit. Psychol.*39, 4 (2018), 793–810. DOI:https://doi.org/10.1111/pops.12463
- [36] Charles S Taber and Milton Lodge. 2006. Motivated skepticism in the evaluation of political beliefs. *Am. J. Pol. Sci.*50, 3 (2006), 755–769. DOI:https://doi.org/10.1111/j.1540-5907.2006.00214.x
- [37] Charles S Taber and Milton Lodge. 2016. The illusion of choice in democratic politics: The unconscious impact of motivated political reasoning. *Polit. Psychol.*37, (2016), 61–85. DOI:https://doi.org/10.1111/pops.12321
- [38] H. Tajfel and J. Turner. 1986. The social identity theory of intergroup behavior. In *Key Readings in Social Psychology. Political Psychology*, J.T. Jost and J. Sidanius (eds.), Psychology Press, 276–293.
- [39] Ben M. Tappin, Gordon Pennycook, and David G Rand. 2018. Rethinking the link between cognitive sophistication and identity-protective bias in political belief formation. *Preprint*. Retrieved from https://doi.org/10.31234/osf.io/yuzfj
- [40] J. Turner, Penelope J. Oakes, S. Alexander Haslam, and Craig McGarty. 1994. Self and collective: Cognition and social context. *Soc. Personal. Soc. Psychol.*20, 5 (1994), 454–469.
- [41] Kerrie L. Unsworth and Kelly S. Fielding. 2014. It's political: How the salience of one's political identity changes climate change beliefs and policy support. *Glob. Environ. Chang.*27, 1 (2014), 131–137. DOI:https://doi.org/10.1016/j.gloenvcha.2014.05.002
- [42] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science (80-)*.359, 6380 (2018), 1146–1151. DOI:https://doi.org/DOI: 10.1126/science.aap9559
- [43] Samuel C. Woolley. 2016. Automating power: Social bot interference in global politics. *First Monday* 21, 4 (2016), 1–24.

A APPENDIX

A.1 Hypothesis for study 1 and their respective results.

	Hypotheses	Confirmed?
H1.1	Incoming information threatens the identity of individuals. As a result, the information is evaluated more poorly than compared to a group whose identity is not threatened by the information.	Yes
H1.2	Identity threat is accompanied by the emotional reactions of anger or anxieties.	Yes
H2.1	By changing the salient identity of individuals to an unthreatened identity, the individuals will evaluate the factual information better.	No
H2.2	An identity salience manipulation for an unthreatened identity will result in no changes in evaluation.	Yes
H3.1	The change from a threatened to an unthreatened identity will also be reflected in a change of experienced emotion. Individuals will experience less anger or anxiety.	Yes
H3.2	The identity salience manipulation for unthreatened identities will result in no changes in emotion experienced.	Yes

A.2 Hypothesis for study 2 and their respective results.

	Hypotheses	Confirmed?
H4.1	If incoming information (misinformation) supports the identity of individuals, the information evaluation is better than compared to a group whose identity is not supported by the information.	Yes
H4.2	This identity support is accompanied by higher levels of enthusiasm.	Yes
H5.1	Changing the salient identity of individuals to a non-supporting identity, the individuals will evaluate the misinformation worse.	No
H5.2	An identity salience manipulation from one non-supporting identity to another will result in no changes in evaluation	Yes
H6.1	The change from a supporting to a non-supporting identity will also be reflected in a change of experienced emotion. Individuals will experience less enthusiasm.	No
H6.2	The identity salience manipulation from nonsupported to another non-supported identity will result in no changes in experienced emotion	Yes