

# How would I be perceived if I challenge individuals sharing misinformation? Exploring misperceptions in the UK and Arab samples and the potential for the social norms approach

Selin Gurgun<sup>1</sup>[0000-0001-9803-955X], Muaadh Noman<sup>2</sup>[0000-0002-5430-8591], Emily Arden-Close<sup>1</sup>[0000-0002-5954-2598], Keith Phalp<sup>1</sup>[0000-0003-4458-686X] and Raian Ali<sup>2</sup>[0000-0002-5285-7829]

<sup>1</sup> Faculty of Science and Technology, Bournemouth University, Poole, UK  
{sgurgun, eardenclose, kphalp}@bournemouth.ac.uk

<sup>2</sup> College of Science and Engineering, Hamad Bin Khalifa University, Doha, Qatar  
{muno32967, raali2}@hbku.edu.qa

**Abstract.** Research conducted in the UK explored the presence of misperceptions, revealing that people anticipated more negative consequences for challenging misinformation on social media. These misperceptions include the anticipation of harming relationships, causing embarrassment and offense to others, the belief that challenging may not yield success and the perception that such behaviour is unacceptable. As the UK culture is characterised as individualistic, we replicated this investigation in a collectivistic culture- Arab societies. Our aim is to explore the differences and similarities of these misperceptions across cultures and to examine whether applying the social norms approach can be a solution to address the inaction towards challenging misinformation. Comparing the UK (N=250) and Arabs (N=212), we showed that, in both cultures there are misperceptions towards challenging misinformation. While misperceptions regarding concerning relationship costs and futility remain consistent across cultures the concerns about causing harm to others and the acceptability of the behaviour differ. Participants in the UK show a higher concern about offense or embarrassment, in contrast, participants in Arab countries exhibit higher misperceptions about injunctive norms, perceiving challenging misinformation as less socially acceptable than it actually is. This study also shows that participants' likelihood to challenge misinformation is influenced by their misperceptions about potential harm to others and perceived injunctive norms. These findings present an opportunity to apply the social norms approach to behaviour change by addressing these misperceptions. Messages emphasising social acceptability of correcting misinformation and highlighting that people appreciate being corrected could serve as powerful tools to encourage users to challenge misinformation.

**Keywords:** Misinformation, Challenging Misinformation, Social Norms, User Corrections, Behaviour Change.

## 1 Introduction

As misinformation spreads faster than genuine news (Vosoughi et al. 2018) the need to combat misinformation has become an important topic. Various solutions have been proposed to combat the spread of misinformation. While most solutions focus on

technologies such as automated detection and correction (Garrett and Weeks 2013; Tanaka et al. 2013; Guo et al. 2022), blockchain technologies (Paul et al. 2019) and natural language processing (de Oliveira et al. 2021; Nakov et al. 2021), these solutions often view users as passive recipients rather than active participants (Fernandez and Alani 2018). However, users' detection and correction can also be an effective way to combat misinformation (Vraga and Bode 2017; Walter and Murphy 2018) and can be as effective as algorithmic corrections (Bode and Vraga 2018). Despite this, people generally do not challenge misinformation they encounter on social media (Chadwick and Vaccari 2019; Tandoc et al. 2020; Tully et al. 2020; Vicol 2020; Bode and Vraga 2021a; Bode and Vraga 2021b)

User behaviour on social media is complex and can be influenced by numerous factors with regard to challenging misinformation, such as diversity in individuals' cognitive processes (Swire and Ecker 2018) and users' concerns regarding their image on social media or their relationship with others (Gurgun et al. 2023b).

Social norms, often described as informal rules and standards guiding social behaviours (Cialdini and Trost 1998), are linked to perceived social pressures that can encourage or discourage specific behaviour (Ajzen 1991). These norms are typically categorised into two dimensions: a) what is common (descriptive norms) and b) what is approved or disapproved (injunctive norms) (Dempsey et al. 2018). In both cases norms inform people's anticipations of others' behaviours and promote conformity. Through the lens of these established norms, people perceive and judge the actions and attitudes of others (Berkowitz 2004). Understanding the user perceptions on challenging behaviour on social media and addressing potential misperceptions allows for correcting them which is an important step to ensure that the decisions related to challenging misinformation are based on reality, rather than misperceptions.

In prior research, scholars have explored the presence of misperceptions towards challenging misinformation on social media (Gurgun et al. 2023a). Their findings indicated that people on social media overestimate the negative reactions they would face when challenging others, a perception that does not align with the actual reactions. However, it is important to note that this sample chosen was drawn from the UK, characterised by a high level of individualism (Hofstede 2023) which may potentially diminish the adherence to social norms (Heinrichs et al. 2006; Schreier et al. 2010) and therefore influence the decision to challenge. Consequently, these results may or may not be applicable in other cultural contexts, such as Arab culture, characterised by high collectivism where high level of empathy and the concerns about preservation of relationships (Hofstede 1984, 2023) may also pose challenges to correcting misinformation. A prevalent critique asserts that psychological studies often rely on WEIRD samples predominantly sourced from Western, Educated, Industrialized, Rich, and Democratic backgrounds (Henrich et al. 2010). To overcome this critique, we collected data from Arabs.

This study is a replication of the prior research that investigates the misperceptions in the UK to Arab participants to explore the differences and similarities of these misperceptions across cultures. The motivation of this work is to pave the way for effective strategies, such as the social norms approach, to mitigate the dissemination of misinformation across diverse cultural contexts.

In the U.K. one of the main reasons preventing people from challenging misinformation is adherence to the norm of conflict avoidance (Chadwick et al. 2022). Research suggests that Western cultures, characterised by high individualism (Hofstede 1984) value open discussion and direct confrontation, whereas collectivistic cultures like Arab culture use avoidance as a conflict management strategy (Adair et al. 2001; Tjosvold and Sun 2002). Therefore, a comparative analysis of these cultures can provide valuable insights into variations in misperceptions between the two cultures and their potential impact on people's likelihood to challenge misinformation.

Drawing from the prior research we used three potential misperceptions that were identified before. The first is "relationship consequences", where users hold negative assumptions about how challenging others might affect their interpersonal relationships. This is based on the idea that confrontation might lead to tension or conflict, potentially negatively impacting relationships (Mutz 2002; Sleeper et al. 2013). Another aspect involves assumptions users hold regarding the potential harm people experience when they are challenged. This includes the risk of people feeling offended or embarrassed when their ideas are contradicted (Steen-Johnsen and Enjolras 2016; Tandoc et al. 2020), which we categorised as "negative impact". In addition, users' beliefs regarding the effectiveness of challenging misinformation represent another misperception, which we named "futility". This conveys a sense of hopelessness or ineffectiveness people may experience with regard to challenging misinformation (Brookover et al. 1978; Van Houtte and Stevens 2008). Lastly, we investigate the acceptability of challenging behaviour referred to as "injunctive norms" (Rimal and Real 2003; Berkowitz 2004; Lapinski and Rimal 2005). We aim to gain insights into whether users' perceptions of challenging misinformation differ across various cultures.

Identifying perceived norms and discrepancies between perceived and actual norms are the first steps in utilising social norms approach (Perkins 2003). Social norms approach is based on the idea of correcting these misperceptions to alleviate the social pressure associated with conforming incorrectly perceived as the norm. Consequently, in this study we also examine whether applying the social norms approach can serve as a solution to address the inaction towards challenging misinformation.

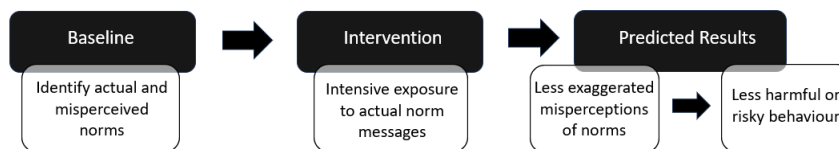
Our contribution lies in replicating and extending prior research conducted in the UK, characterised by individualism, to Arab societies characterised by collectivism. By comparing both cultures, we aim to explore whether the observed misperceptions in the UK are a culturally specific phenomenon or if they occur similarly in a more collectivistic culture. This provides insights into the interaction between cultural norms and individuals' likelihood to challenge misinformation on social media. Furthermore, this research investigates the potential of the social norms approach as a cross-cultural solution to address inaction towards challenging misinformation. Ultimately, our work aims to foster an environment conducive to making more informed decisions when encountering misinformation.

## 2 Theoretical Underpinnings

### 2.1 Social Norms Approach

Social norms theory refers to situations where people hold incorrect perception of the attitudes and behaviours of others. This theory has been applied in health related areas (Dempsey et al. 2018) particularly to peer substance use (Perkins 2003; McAlaney and McMahon 2007; McAlaney et al. 2015) where the majority tends to overestimate the frequency and amount of substance use and its acceptability among their peers, leading to an increase in their own consumption. This misperception is identified by comparing their own behaviour and their perception of the norm held by the majority of their peers (McAlaney et al. 2015). Such perceptions act like social influences on people, increasing their consumption to conform the estimated norm (Borsari and Carey 2003). In short, individuals behaviour is significantly shaped by their assumption of what is considered “normal”(Bandura 1986).

In terms of prevention and intervention of substance use, social norms approach involves challenging students’ misperceptions about peer substance use and attitudes (Perkins et al. 2003; McAlaney et al. 2015). The foundation of these interventions is to highlight the actual norm, which is often healthier and safer than the perceived norm. As an example, if students think that their peers typically consume six alcoholic drinks at parties, when in reality, they only have three, a social norms campaign might give the message that the majority of students typically have no more than three alcoholic drinks when they party. The idea is that correcting these misperceptions reduces the social pressure to follow what is incorrectly perceived as the norm, leading to reduced e.g., alcohol use and more negative attitudes toward it (Perkins et al. 2003) (See Fig. 1).



**Fig. 1.** Model of Social Norms Approach to Prevention (Perkins 2003)

Social norms theory can also be applied to different situations where individuals avoid confronting others’ problem behaviour due to their misperceived belief that such behaviour is socially acceptable within their peer group (Berkowitz 2003). In other words, people who underestimate how much their peers are uncomfortable with problematic behaviour might choose not to express their discomfort regarding that behaviour. However, if people were made aware of the actual level discomfort their peers feel about that behaviour, they might become more willing to voice their concerns. For instance, research in the U.S. indicates that most college students tend to underestimate how accepting and supportive their peers are of gay, lesbian and bisexual students. They believe their peers are less tolerant and supportive than they actually are (Bowen and Bourgeois 2001). Another study in the U.S. found that, college students tend to

misperceive the risky sexual behaviour norms of their peers, believing that their peers engage in riskier sexual behaviour than they actually do. These misperceptions have been positively associated with students' sexual behaviour (Lewis et al. 2007). Social norms theory has also been applied in the context of distracted driving behaviour. Adolescents tend to perceive that their parents and peers engage in distracted driving more frequently than they do and this misperception can lead them to underestimate the risks associated with distracted driving, leading to their engagement in such behaviour (Carter et al. 2014).

Considering the influence of misperceptions on shaping behaviour, exploring their impact on challenging misinformation can provide valuable insights. It stands to reason that people may be hesitant to challenge misinformation when they encounter it. Expanding upon prior research, this study aims to provide insights into informed interventions and strategies that transcend specific cultural contexts. Since the social norms approach has effectively influenced a wide range of behaviors, this research expanded its application to investigate how users perceive the act of challenging misinformation with a particular emphasis on the examination of cultural differences.

### **3 Method**

#### **3.1 Participants**

A total of 462 participants (age range 18 – 77 years, 176 female, 283 male and 3 non-binary, 250 from the U.K. and 212 from Arab countries (Egypt 22.2%, Saudi Arabia 12.3%, Iraq 10.8%, Jordan 9.9%, Syria 9%, Bahrain 9%, Lebanon 7.5%, Oman 6.6%, Palestine 4.2%, Kuwait 3.3%, Algeria 1.4%, Morocco 1.4%, Tunisia 0.9%, Sudan 0.5%, Yemen 0.5%, and the United Arab Emirates 0.5%) were recruited through Prolific<sup>TM</sup> ([www.prolific.co](http://www.prolific.co)) and Cint ([www.cint.com](http://www.cint.com)), well-established online recruitment platforms for research studies. Inclusion criteria were being 18 years or older, fluency in the English or Arabic language, using Facebook with their identity, prior exposure to misinformation on Facebook, and being based in the UK or an Arabic country-based.

#### **3.2 Data Collection**

The survey was conducted online using Qualtrics, an online survey design platform. Participants were informed about the study objectives and asked for their consent before proceeding. They were given information regarding data confidentiality, their freedom to participate and the right to withdraw from the study and their access to the study findings. The ethics approval was obtained by Bournemouth University and Hamad Bin Khalifa University. The questionnaire included three attention checks, and participants who failed at least two of them were excluded from the analysis. In recognition of their participation, eligible participants were compensated for their participation.

### 3.3 Questionnaire Design

The questionnaire for this study was created using Qualtrics™ (<https://www.qualtrics.com>), an online survey platform. To ensure participants had a clear and common understanding, we included the key concepts' definitions such as "misinformation", "challenging" and "acquaintance". The questions asking challenging acquaintances was intentional as we tried to maintain a neutral stance between total strangers and loved ones in social relationships, since people's behaviour can vary when interacting with strangers and loved ones (Tandoc et al. 2020). Facebook was chosen due to its global base with approximately 2.91 billion monthly active users (Statista 2016) and being the most frequent mentioned (67%) to be a platform where users encountered fake news (Centre for International Governance and Ipsos 2019). It is also an appropriate platform for investigating social norms as it is a semi-public space where comments and likes can be seen by others apart from friends. Facebook also consists of both strong ties (e.g., family) and weak ties (e.g., acquaintances like neighbours) (Ellison and Boyd 2013). Therefore, it is a convenient platform to investigate the attitudes towards acquaintances. Weak ties are significant as people are more likely to encounter different opinions through such connections (Kim 2011) people they find it more difficult to confront them than to confront those with whom they have strong ties (Valenzuela et al. 2012).

The survey evaluated Arab and UK participants' attitudes and their perceptions of others regarding challenging misinformation on social media. Participants were first asked about their attitudes towards challenging misinformation which we term "self-term" in this paper and after they were asked about their opinions about others which we term "perceived". The questionnaire along with the data can be found in OSF (<https://osf.io/p2yav/>).

### 3.4 Measures

#### **Demographic Characteristics.**

Participants answered questions about their age, gender and educational level.

#### **Relationship consequences.**

We utilised the items from a novel scale of Perceived Relationship Costs (Zhang et al. 2011; Zhang and Wei 2017). Example of self-reported items included are "*It would offend me*" and "*I would think that they are aggressive*". Internal reliability was high for both UK ( $\alpha = 0.92$ ) and Arabs ( $\alpha = 0.91$ ). For the perceived items, we changed the items as "*It would offend them*" and "*They would think that I am aggressive*". Internal reliability was also high for both UK ( $\alpha = 0.92$ ) and Arabs ( $\alpha = 0.93$ ).

#### **Negative impact on the person being challenged.**

Self-reported negative impact and perceived negative impact were assessed with two items on a 7-point Likert scale from Strongly Disagree to Strongly Agree. The items are for self-report "*I would feel embarrassed*" and "*I would feel that they will be viewed as untrustworthy by other social contacts*". The items for perceived are "*They would*

*feel embarrassed*” and *“They would feel that I will be viewed as untrustworthy by other social contacts”*. The items were derived from previous research (Milliken et al. 2003; Tandoc et al. 2020; Altay et al. 2022). Given the acknowledged limitations associated with using Cronbach’s alpha for two-item scales, we did not include it in our results (Hulin and Cudeck 2001).

#### **Futility of the act of challenging.**

Actual futility and perceived futility were assessed with two items on a 7-point Likert scale from Strongly Disagree to Strongly Agree adapted from Milliken et al.’s (2003) study on employee silence. Self-reported items include *“challenging would change my mind that the information I shared is true”* and for perceived *“challenging would change their mind that the information they shared is true”*.

#### **Injunctive norms.**

Participants were asked to respond to the self-report item, *“how do you find challenging others on Facebook when they share misinformation?”* and the perceived item *“how would a typical person from your Facebook network find challenging others when they share misinformation?”*. The responses were recorded on a 7-point scale from (1) very unacceptable to (7) very acceptable.

### **3.5 Data Analysis**

Data were analysed using SPSS version 28. Descriptive statistics were used to report demographic information and the range, mean, and standard deviation for the misperception variable in both Arab and UK populations. As both Arab and UK data was not normally distributed, non-parametric tests were used. A Mann-Whitney U test was run to determine if there were differences in misperception scores between the UK and Arab countries. Misperception score was calculated for each participant by subtracting the self-report mean from the perception of the participants (them score). A result of a positive misperception score was interpreted as an overestimation (further details on the method can be found in (Perkins et al. 2010; Kenney et al. 2013; Sandstrom et al. 2013). Ordinal logistic regression was used to assess the impact of the misperceptions on the likelihood to challenge.

## **4 Results**

### **4.1 Participant Demographics**

In total, 585 participants (324 from the Arab region, 261 from the UK) completed the online survey. Of those who completed, 462 participants (212 Arab, 250 UK) provided valid answers, passed three attention checks, and not being considered speeders or duplicates. Demographic characteristics of the participants, including age, gender, and educational background, are presented in Table 1.

**Table 1.** Participants demographics

		Arab (N=212)		UK (N=250)	
		Count	%	Count	%
Gender	Male	140	66.0	104	41.6
	Female	72	34.0	143	57.2
	Non-binary	0	0	3	1.2
Age	18-24	31	14.6	44	17.6
	25-34	113	53.3	94	37.6
	35-44	50	23.6	67	26.8
	Over 45	18	8.5	45	18.0
Education	Primary	14	6.6	36	14.4
	Further	29	13.7	57	26.8
	Higher	169	79.7	157	62.8

#### 4.2 Misperceptions of relationship costs, negative impact on others, futility, and injunctive norms

A paired-samples t-test was used to examine the discrepancy between participants' actual attitudes and their perceptions of others' attitudes towards being challenged when posting misinformation. The significant difference between their self-report and perceived responses indicates that there is a misperception.

During data screening, only four outliers were identified for the UK and one outlier was identified for the Arabs each falling more than 1.5 box-lengths from the edge of the box in a box plot. However, their values did not classify as extreme, so they were kept in analyses. For both data assumption of normality was not violated, as assessed by a normal Q-Q plot for all the analysis. Inspection of their values did not reveal them to be extreme and they were kept in the analysis. The assumption of normality was not violated, as assessed by a normal Q-Q plot for all the analysis.

Table 2 shows the statistically significant differences in Arab participants for all six perceived relationship cost items ( $p < .01$ ). This is in line with the UK findings (Gurgun et al. 2023a). Arab participants reported greater anticipation of other's negative reactions regarding their relationships compared to their actual reactions when being challenged. The mean of self-report for the relationship costs was 3.09 (SD = 1.57) and for perceived relationship costs, it is 3.70 (SD = 1.53) with a significant t-value of -6.11 ( $p < .001$ ). This suggests that Arab participants perceive challenging others as more detrimental to their relationships compared to how they would react if they were challenged.

In terms of the items relating to negative impact on others, the misperceptions for the negative impact was not statistically significant for either cultural group.

For the items around futility of challenging misinformation, the analysis involved reverse-coding positively framed items. For Arab participants, the results indicated a significant difference for all the items ( $p < .001$ ) in the Arab sample whereas in the UK only one item "It would make me (them) delete the post" was significant ( $p < .01$ ). The mean of self-report for futility was 2.59 (SD = 1.43) and for perception of others it was 3.12 (SD = 1.22), with a significant t-value of -4.65 ( $p < .001$ ). This suggests that Arab

#### Exploring misperceptions in UK and Arab samples

participants think that their effort will be futile as others are unlikely to change their mind and delete their post; however they themselves are more inclined to change their mind and delete their post.

Similar to the UK results, among Arab participants there was a significant difference between participants' self-reported answers and their perception of others answers regarding the acceptability of challenging misinformation ( $p < .001$ ). The mean of self-report for injunctive norms was 4.79 ( $SD = 1.80$ ) compared to 4.4 ( $SD = 1.55$ ) for perception of others, with a significant t-value of 2.71 ( $p < .01$ ). This indicates that, participants perceive that others consider it less acceptable to challenge misinformation than they find it to be.

**Table 2.** Mean and SD values of self-reported and perceived items across UK and Arab samples

	Self-Report		Perceptions		t-value
	Arab (N = 212) M (SD)	U.K (N = 250) M (SD)	Arab (N = 212) M (SD)	U.K (N = 250) M (SD)	
Relationship Consequences					
It would offend me (them)	3.51 (1.96)	3.78 (1.83)	4.04 (1.79)	4.72 (1.46)	-3.75***
I (They) would think that they are (I am) aggressive	2.87 (1.72)	3.64 (1.71)	3.76 (1.83)	4.36 (1.57)	-6.63***
I (They) would think that they are (I am) unfriendly	3.07 (1.83)	3.66 (1.7)	3.76 (1.88)	4.48 (1.52)	-5.13***
I (They) would think that they are (I am) not empathetic	3.04 (1.87)	3.44 (1.67)	3.7 (1.87)	4.3 (1.47)	-4.97***
The relationship between us will deteriorate	2.74 (1.84)	3.58 (1.77)	3.25 (1.73)	4.26 (1.55)	-4.17***
We will interact less frequently afterwards	3.3 (1.9)	4.11 (1.79)	3.68 (1.76)	4.52 (1.52)	-3.12**
Average perceived relationship costs	3.09 (1.57)	3.71 (1.49)	3.7 (1.53)	4.43 (1.29)	-6.11***
Negative impact on the person being challenged					
I (They) would feel embarrassed or upset	4.19 (2.02)	4.86 (1.58)	4.44 (1.69)	4.79 (1.63)	-2.04*
I (They) would feel that I will be viewed as untrustworthy by other social contacts	3.9 (2.01)	4.62 (1.65)	4.02 (1.73)	4.4 (1.44)	-0.97
Average perceived negative items	4.04 (1.87)	4.74 (1.48)	4.23 (1.5)	4.6 (1.27)	-1.77
Futility					
It would change my (their) mind (R)	2.67 (1.72)	3.6 (1.47)	3.2 (1.48)	3.88 (1.36)	-3.79***
It would make me (them) delete the post (R)	2.52 (1.67)	3.31 (1.74)	3.04 (1.44)	4.01 (1.45)	-3.86***
Average perceived futility items	2.59 (1.43)	3.46 (1.40)	3.12 (1.22)	3.95 (1.21)	-4.65***
Injunctive Norms					
How do you (would a typical person) find challenging others on Facebook when they share misinformation?	4.74 (1.8)	4.95 (1.42)	4.4 (1.55)	4.29 (1.34)	2.71**
					6.83***

\* p &lt; 0.05, \*\* p &lt; 0.01, \*\*\* p &lt; 0.001

A Mann-Whitney U test was run to determine any differences in degree of misperception between UK and Arab countries. In line with prior research on misperceptions (Perkins et al. 2010; Kenney et al. 2013; Duong and Parker 2018), we computed misperception scores as the perceived (“them” score) minus average self-reported scores (average of all “me” scores). Misperception scores for UK participants ranged from -3.95 to 3.69 and for Arab participants ranged from -3.74 to 4. Distributions of the scores for UK and Arab countries were similar, as assessed by visual inspection. There was no statistically significant difference between Arabs (median 0.59) and UK (median 0.97) regarding the misperception of relationship costs score,  $U = 23874$ ,  $z = -1.84$ ,  $p = 0.07$ . Similarly, there was no statistically significant difference between Arabs (0.41) and UK (0.54) regarding the misperception of futility,  $U = 25244$ ,  $z = -0.88$ ,  $p = 0.38$ . For the misperception of negative impact on others, median was statistically significantly higher in the U.K (0.08) than in Arab countries (-0.42),  $U = 22185$ ,  $z = -3.04$ ,  $p = 0.002$ . The results indicates that, participants in Arab countries tend to underestimate the negative impact on others when challenged, perceiving that they themselves will experience more negative harm than their perception of others. In contrast, in the UK, participants are in an overestimation, thinking that others will experience more negative harm from them than they actually do. However, it is also worth noting that the effect size was 0.14 which is a small effect based on Cohen’s classification (Cohen 2013). For the injunctive norms, median was statistically significantly lower in the UK (-0.95) than in Arab countries (-0.74),  $U = 33136$ ,  $z = 4.67$ ,  $p < 0.001$ . This indicates that, in the Arab countries there is a greater misperception that challenging others is unacceptable compared to the UK. The effect size was 0.22 which is small to medium effect size category.

**Table 3** Comparisons of misperception results by the Mann-Whitney U test

Variables	Median	Median	U	z	p
Misperception relationship costs	0.97	0.59	23874	-1.84	0.07
Misperception futility	0.54	0.41	25244	-0.88	0.38
Misperception negative impact on others	0.08	-0.42	22185	-3.04	0.002
Misperception injunctive norms	-0.95	-0.74	33136	4.67	<0.001

### 4.3 The affect of misperception on the likelihood to challenge

For the purposes of this study, prior to the regression analysis, we treated misperception data as overestimation as on average, participants are more likely to overestimate the negative reactions from other when they challenge. To operationalise the data, we categorised the misperception scores. A positive misperception score (i.e., actual attitude of participants < perceived attitude of others) indicated a tendency to overestimate others’ negative attitude in terms of challenging misinformation. We classified the negative misperception score (i.e., actual attitude of participants > perceived attitude

of others) and zero score (i.e., actual attitude of participants = perceived attitude of others) as having no overestimation. This approach allowed for grouping of participants based on whether they overestimated.

A cumulative odds ordinal logistic regression with proportional odds was run to determine the extent to which overestimation of others' negative reactions impacts likelihood to challenge across the two cultures, controlling for age and gender (See Table X)

For the Arab countries, the assumption of proportional odds was met, as assessed by a full likelihood ratio test comparing the fit of the proportional odds model to a model with varying location parameters,  $\chi^2(210) = 232.484$ ,  $p = .137$ . However, for the U.K, this assumption was not met,  $\chi^2(255) = 676.862$ ,  $p < .001$ . Therefore, an examination of the assumption of proportional odds was undertaken by running separate binomial logistic regressions on cumulative dichotomous dependent variables. This examination showed that for most of the variables the assumption of proportional odds appears tenable. The deviance goodness-of-fit test indicated that for both cultures, the model was a good fit to the observed data, for the UK  $\chi^2(1203) = 736.312$ ,  $p = 1$  and Arab countries  $\chi^2(1050) = 661.438$ ,  $p = 1$ . For both cultures, the final model statistically significantly predicted the dependent variable over and above the intercept-only model, UK  $\chi^2(51) = 99.733$ ,  $p < .001$  and for Arabs  $\chi^2(42) = 61.494$ ,  $p = .026$ . Regarding the independent variables, for the UK age, overestimating negative impact and overestimating injunctive norms significantly predicted likelihood to challenge (odds ratio of 1.034 (95% CI, 1.013 to 1.056)  $p < .01$ , odds ratio of 0.585 (95% CI, 0.352 to 0.971),  $p < .05$ . and odds ratio of 1.581 (95% CI, 1.001 to 2.498),  $p = .05$  respectively). For the Arab countries, only overestimating injunctive norms significantly predicted likelihood to challenge (odds ratio of 1.790 (95% CI, 1.085 to 2.951)  $p < .05$ )

**Table 7.** Ordinal logistic regression results for predicting the likelihood to challenge with predictors including age, gender, overestimating the three negative consequences and injunctive norms

Predictors	UK			Arab Countries		
	B	SE	Odds Ratio	B	SE	Odds Ratio
Age	0.034	0.01	1.034*	0.011	0.015	1.011
Gender (Female)	-0.38	0.231	0.684	-0.287	0.261	0.75
Overestimating Relationship Consequences	-0.063	0.285	0.939	-0.451	0.288	0.637
Overestimating Futility	0.166	0.241	1.181	0.107	0.268	1.113
Overestimating Negative Impact (offending others)	-0.536	0.259	0.585*	-0.198	0.285	0.82
Overestimating Injunctive Norms (acceptability)	0.458	0.233	1.581*	0.582	0.255	1.790*

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

## 5 Discussion

Our findings highlight the presence of misperceptions in both cultural contexts, highlighting the potential for Social Norms Approach to address and correct these misperceptions. Prior research showed the existence of misperceptions across these variables in the UK (Gurgun et al. 2023a). The new results from Arab participants are aligned with those from the UK, indicating the existence of these misperceptions. In both cultures, participants believed that, through challenging misinformation, their relationship with others will be harmed and people will be offended or embarrassed. They also perceive that challenging misinformation is not socially acceptable. Our study showed that these perceptions do not align with reality. Based on the difference between their self-report and perceptions, we can conclude that, in both cultures, participants overestimate the potential harm of challenging misinformation to their relationships, the negative impact on others and the lack of potential success while underestimating the acceptability of this behaviour. Recognising these misperceptions is important as it allows for their correction, in turn leading to an environment where people are more motivated to engage in correcting misinformation. This provides an opportunity for the social norms approach, which aims to correct misperceptions of norms and promote more positive behaviours. The social norms approach predicts that interventions to correct misperceptions by revealing the actual, beneficial norm will have a positive impact on individuals, leading to a decrease in problem behaviours (Perkins 2003; Berkowitz 2004). Therefore, by addressing misperceptions and promoting accurate norms, the social norms approach can be an effective strategy for behaviour change in both individualistic and collectivist cultures.

The analysis of differences in misperception scores between the UK and Arab cultures revealed that participants from both cultures have the tendency to overestimate or underestimate the negative consequences and social norms associated with challenging misinformation.

While some misperceptions did not differ across cultures, such as misperception of relationship costs and futility; negative impact on others and injunctive norms were significantly different across both cultures. Cultural values and tendencies of the participants towards confrontation and conflict would provide further context to understand these similarities and differences.

The concern regarding causing harm to others, including offense or embarrassment when challenging is higher in the UK. This might be due to differences in communication styles. According to Hofstede's dimensions (Hofstede 2023), the UK has been defined as an individualistic culture which places a high value on open communication and direct confrontation which may manifest in a more assertive approach to challenging misinformation (Morris et al. 1998; Tjosvold and Sun 2002). This approach, however, in contrast with the politeness norm in the UK which is characterised by values like respect, respecting privacy, keeping distance and being reserved (Kamehkhosh and Larina 2020). This communication style is often perceived as avoidance-based and

hearer oriented (Larina 2008) which promote indirectness and subtlety in communication. While individuals in the UK may prioritise open communication, they may also be mindful of social norms and have concerns more about others when they challenge and aim to avoid causing embarrassment or offense when correcting.

Participants in Arab countries displayed higher misperceptions about injunctive norms, indicating they believe that challenging misinformation is less socially acceptable than it actually is. In collectivistic cultures such as Arab culture, there is a greater emphasis on maintaining social harmony and indirect communication (Triandis 1994; Tjosvold and Sun 2002; Zhang et al. 2011). One key feature that defines the people in collectivist cultures is their strong emphasis on relationships (Zhang et al. 2011). Challenging misinformation therefore can be perceived less acceptable due to the cultural norm of preserving the existing interpersonal relationships. Another aspect to consider might be the role of power distance which refers to the extent to which less powerful members of society accept and expect that power is distributed unequally (Hofstede 1984). In hierarchical societies with high power distance such as many Arab countries, challenging misinformation can be perceived as less acceptable, especially if misinformation is propagated by authority figures or those perceived to hold higher status, due to respect for hierarchical structures.

Our results demonstrate that certain misperceptions have an impact on likelihood of challenging misinformation, in line with previous research highlighting the significant impact of misperceptions on behaviour and individuals' decision to engage in specific behaviour (McAlaney et al. 2020).

In the UK participants who overestimated the negative impact on others of challenging misinformation were less inclined to challenge misinformation. The concern for the potential harm on others significantly influenced their decision to challenge. Additionally, in both cultural contexts, perceived injunctive norms, the acceptability of the behaviour, play an important role in likelihood of challenging misinformation. When people perceive challenging others is socially acceptable, they are more inclined to challenge misinformation.

These findings pave the way for tailoring interventions and strategies to address these misperceptions. In addition to its success in various domains such as health related behaviour (Reid and Aiken 2013; Dempsey et al. 2018), substance use (Perkins 2003; McAlaney et al. 2015) and smartphone use (McAlaney et al. 2020), the social norms approach can be applied to different situations where individuals refrain from a behaviour due to their misperceived belief that such behaviour is not socially acceptable within their peer group (Berkowitz 2003). In our study, although many participants reported that they find challenging misinformation is socially acceptable, they believe that others do not consider it socially acceptable. However, research revealed that people are reluctant to share misinformation due to fear of potential damage to their reputation that they are only willing to do so in exchange for a monetary incentive (Altay et al. 2022). In this case, this contradiction suggests, people might value being challenged as it can help protect their own reputation as opposed to participants' perception.

When implementing the social norms approach, messages that convey the idea that challenging misinformation is socially acceptable and that people appreciate being corrected, hold the premise to motivate people to challenge misinformation more

confidently. Through this approach, an environment where challenging misinformation is not seen as personal and recognised as common can be cultivated. This can help alleviate misperceptions such as fear of causing offense or the belief that it is socially unacceptable.

### **5.1 Threats to validity**

There are potential threats to validity that could impact the quality and generalisability of our findings. First, when asking people about their behaviour, one must consider the Hawthorne effect (Wickström and Bendix 2000) which refers to the situations where individuals change their behaviour when they are aware of being observed. The discrepancies could also be due to self-enhancement biases such as illusory superiority where people tend to overstate their positive characteristics and underestimate their negative characteristics compared to the average other or most others (Brown 1986; Hoorens 1993). Self-reports of experience might be subject to recall error (Sudman et al. 1996) or social desirability bias (Grimm 2010). To address these, future work should consider an experimental study or observational research to gain a better understanding of participants' behaviour. We selected specific cultures such as the UK and Arab countries as representations of individualistic and collectivistic cultures. The results might not be generalisable to other cultures with similar or different characteristics. To address this limitation, future work should consider comparative studies involving a broader range of cultures.

## **6 Conclusion and Future Work**

Our study reveals misperceptions about challenging misinformation in both UK and Arab cultural contexts. Despite cultural differences, participants in both settings tend to overestimate harm to relationships and underestimate the social acceptability and effectiveness of challenging misinformation. The difference between self-reported and actual norms is crucial for motivating challenging behaviour. The social norms approach emerges as a promising strategy to address these misperceptions, correcting norms and promoting positive behaviours. Although the differences in misperception scores between the UK and Arab cultures highlight cultural nuance, interventions based on the social norms approach, revealing the healthy norm, can positively impact individuals in both cultures, fostering a reduction in problem behaviours. Our findings have broader implications extending the application of the social norms approach to online interactions where interventions should emphasise the social acceptability of the behaviour. Additional to self-reports, future work could consider observational data to increase the validity of the findings. Moreover, this study's findings can be applied to understand whether the misperceptions identified are common across various cultures beyond the UK and Arab context. Additionally, future studies might explore the implementation of the social norms approach in the design of social media platforms to assess its potential positive impact on the frequency of challenging misinformation in online sphere.

**Acknowledgments.** This publication was supported by NPRP 14 Cluster grant # NPRP 14C-0916–210015 from the Qatar National Research Fund (a member of Qatar Foundation). The findings herein reflect the work and are solely the responsibility of the authors.

## References

- Adair, W. L., Okumura, T. and Brett, J. M., 2001. Negotiation behavior when cultures collide: The United States and Japan. *Journal of Applied Psychology*, 86 (3), 371.
- Ajzen, I., 1991. The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50 (2), 179-211.
- Altay, S., Hacquin, A.-S. and Mercier, H., 2022. Why do so few people share fake news? It hurts their reputation. *New Media & Society*, 1461444820969893.
- Bandura, A., 1986. Social foundations of thought and action. *Englewood Cliffs, NJ*, 1986 (23-28).
- Berkowitz, A. D., 2003. Applications of social norms theory to other health and social justice issues. In: Perkins, H. W., ed. *The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians*. Hoboken, NJ: Jossey-Bass/Wiley, 259-279.
- Berkowitz, A. D., 2004. The social norms approach: Theory, research, and annotated bibliography: Citeseer.
- Bode, L. and Vraga, E. K., 2018. See Something, Say Something: Correction of Global Health Misinformation on Social Media. *Health Communication*, 1131.
- Bode, L. and Vraga, E. K., 2021a. Correction Experiences on Social Media During COVID-19. *Social Media + Society*, 7.
- Bode, L. and Vraga, E. K., 2021b. Value for Correction: Documenting Perceptions about Peer Correction of Misinformation on Social Media in the Context of COVID-19. *Journal of Quantitative Description: Digital Media*, 1.
- Borsari, B. and Carey, K. B., 2003. Descriptive and injunctive norms in college drinking: a meta-analytic integration. *Journal of studies on alcohol*, 64 (3), 331-341.
- Bowen, A. M. and Bourgeois, M. J., 2001. Attitudes toward lesbian, gay, and bisexual college students: The contribution of pluralistic ignorance, dynamic social impact, and contact theories. *Journal of american college health*, 50 (2), 91-96.
- Brookover, W. B., Schweitzer, J. H., Schneider, J. M., Beady, C. H., Flood, P. K. and Wisenbaker, J. M., 1978. Elementary school social climate and school achievement. *American educational research journal*, 15 (2), 301-318.
- Brown, J. D., 1986. Evaluations of self and others: Self-enhancement biases in social judgments. *Social cognition*, 4 (4), 353-376.
- Carter, P. M., Bingham, C. R., Zakrajsek, J. S., Shope, J. T. and Sayer, T. B., 2014. Social norms and risk perception: Predictors of distracted driving behavior among novice adolescent drivers. *Journal of Adolescent Health*, 54 (5), S32-S41.
- Centre for International Governance, I. and Ipsos, 2019. CIGI-Ipsos Global Survey on Internet Security and Trust.
- Chadwick, A. and Vaccari, C., 2019. *News sharing on UK social media: Misinformation, disinformation, and correction*. Loughborough: Online Civic Culture Centre, Loughborough University.

## Exploring misperceptions in UK and Arab samples

- Chadwick, A., Vaccari, C. and Hall, N.-A., 2022. *Covid vaccines and online personal messaging: the challenge of challenging everyday misinformation*. Loughborough University.
- Cialdini, R. B. and Trost, M. R., 1998. Social influence: Social norms, conformity and compliance. In: Gilbert, D. T., Fiske, S. T. and Lindzey, G., eds. *The handbook of social psychology*, Vols. 1-2, 4th ed. New York, NY: McGraw-Hill, 151-192.
- Cohen, J., 2013. *Statistical power analysis for the behavioral sciences*. Academic press.
- de Oliveira, N. R., Pisa, P. S., Lopez, M. A., de Medeiros, D. S. V. and Mattos, D. M. F., 2021. Identifying fake news on social networks based on natural language processing: trends and challenges. *Information*, 12 (1), 38.
- Dempsey, R. C., McAlaney, J. and Bewick, B. M., 2018. A critical appraisal of the social norms approach as an interventional strategy for health-related behavior and attitude change. *Frontiers in psychology*, 9, 2180.
- Duong, H. T. and Parker, L., 2018. Going with the flow: Young motorcyclists' misperceived norms and motorcycle speeding behaviour. *Journal of Social Marketing*, 8 (3), 314-332.
- Ellison, N. B. and Boyd, D., 2013. Sociality through social network sites. *The Oxford handbook of internet studies*, 151-172.
- Fernandez, M. and Alani, H., 2018. Online misinformation: Challenges and future directions, *Companion Proceedings of the The Web Conference 2018* (pp. 595-602).
- Garrett, R. K. and Weeks, B. E., 2013. The promise and peril of real-time corrections to political misperceptions, *Proceedings of the 2013 conference on Computer supported cooperative work* (pp. 1047-1058).
- Grimm, P., 2010. Social desirability bias. *Wiley international encyclopedia of marketing*.
- Guo, Z., Schlichtkrull, M. and Vlachos, A., 2022. A survey on automated fact-checking. *Transactions of the Association for Computational Linguistics*, 10, 178-206.
- Gurgun, S., Cemiloglu, D., Arden-Close, E., Phalp, K., Ali, R. and Nakov, P., 2023a. Challenging Misinformation on Social Media: Users' Perceptions and Misperceptions and Their Impact on the Likelihood to Challenge, *Available at SSRN 4600006*.
- Gurgun, S., Cemiloglu, D., Arden-Close, E., Phalp, K., Nakov, P. and Ali, R., 2023b. Why Do We Not Stand Up to Misinformation? Factors Influencing the Likelihood of Challenging Misinformation on Social Media and the Role of Demographics, *Factors Influencing the Likelihood of Challenging Misinformation on Social Media and the Role of Demographics*.
- Heinrichs, N., Rapee, R. M., Alden, L. A., Bögels, S., Hofmann, S. G., Oh, K. J. and Sakano, Y., 2006. Cultural differences in perceived social norms and social anxiety. *Behaviour research and therapy*, 44 (8), 1187-1197.
- Henrich, J., Heine, S. J. and Norenzayan, A., 2010. The weirdest people in the world? *Behavioral and brain sciences*, 33 (2-3), 61-83.
- Hofstede, G., 1984. *Culture's consequences: International differences in work-related values*. Vol. 5. sage.
- Hofstede, G., 2023. *Country Comparison Tool* [online]. <https://www.hofstede-insights.com/country-comparison-tool>: Available from: <https://www.hofstede-insights.com/country-comparison-tool> [Accessed
- Hoorens, V., 1993. Self-enhancement and superiority biases in social comparison. *European review of social psychology*, 4 (1), 113-139.

- Hulin, C. and Cudeck, R., 2001. Cronbach's alpha on two-item scales. *Journal of Consumer Psychology*, 10 (1/2), 55.
- Kamehkhosh, N. and Larina, T. V., 2020. Cultural values and politeness strategies in British and Persian family discourse. *Proceedings of INTCESS*, 2020, 7th.
- Kenney, S. R., LaBrie, J. W. and Lac, A., 2013. Injunctive peer misperceptions and the mediation of self-approval on risk for driving after drinking among college students. *Journal of health communication*, 18 (4), 459-477.
- Kim, Y., 2011. The contribution of social network sites to exposure to political difference: The relationships among SNSs, online political messaging, and exposure to cross-cutting perspectives. *Computers in Human Behavior*, 27 (2), 971-977.
- Lapinski, M. K. and Rimal, R. N., 2005. An explication of social norms. *Communication Theory*, 15 (2), 127-147.
- Larina, T., 2008. Directness, imposition and politeness in English and Russian. *Cambridge ESOL: Research Notes*, 33, 33-38.
- Lewis, M. A., Lee, C. M., Patrick, M. E. and Fossos, N., 2007. Gender-specific normative misperceptions of risky sexual behavior and alcohol-related risky sexual behavior. *Sex Roles*, 57 (1-2), 81-90.
- McAlaney, J., Almourad, M. B., Powell, G. and Ali, R., 2020. Perceptions and misperceptions of smartphone use: Applying the social norms approach. *Information*, 11 (11), 513.
- McAlaney, J., Helmer, S. M., Stock, C., Vriesacker, B., Van Hal, G., Dempsey, R. C., Akvardar, Y., Salonna, F., Kalina, O. and Guillen-Grima, F., 2015. Personal and perceived peer use of and attitudes toward alcohol among university and college students in seven EU countries: Project SNIPE. *Journal of studies on alcohol and drugs*, 76 (3), 430-438.
- McAlaney, J. and McMahon, J., 2007. Normative beliefs, misperceptions, and heavy episodic drinking in a British student sample. *Journal of studies on alcohol and drugs*, 68 (3), 385-392.
- Milliken, F. J., Morrison, E. W. and Hewlin, P. F., 2003. An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of management studies*, 40 (6), 1453-1476.
- Morris, M. W., Williams, K. Y., Leung, K., Larrick, R., Mendoza, M. T., Bhatnagar, D., Li, J., Kondo, M., Luo, J.-L. and Hu, J.-C., 1998. Conflict management style: Accounting for cross-national differences. *Journal of International Business Studies*, 29 (4), 729-747.
- Mutz, D. C., 2002. The consequences of cross-cutting networks for political participation. *American Journal of Political Science*, 838-855.
- Nakov, P., Da San Martino, G., Elsayed, T., Barrón-Cedeño, A., Míguez, R., Shaar, S., Alam, F., Haouari, F., Hasanain, M. and Mansour, W., 2021. Overview of the clef-2021 checkthat! lab on detecting check-worthy claims, previously fact-checked claims, and fake news (pp. 264-291): Springer.
- Paul, S., Joy, J. I., Sarker, S., Ahmed, S. and Das, A. K., 2019. Fake news detection in social media using blockchain, 2019 7th International Conference on Smart Computing & Communications (ICSCC) (pp. 1-5): IEEE.
- Perkins, H. W., 2003. The emergence and evolution of the social norms approach to substance abuse prevention. *The social norms approach to preventing school and college age substance abuse: A handbook for educators, counselors, and clinicians*, 3-17.

## Exploring misperceptions in UK and Arab samples

- Perkins, H. W. J. T. s. n. a. t. p. s., college age substance abuse: A handbook for educators, c. and clinicians, 2003. The emergence and evolution of the social norms approach to substance abuse prevention. 3-17.
- Perkins, J. M., Perkins, H. W. and Craig, D. W., 2010. Peer weight norm misperception as a risk factor for being over and underweight among UK secondary school students. *European Journal of Clinical Nutrition*, 64 (9), 965-971.
- Reid, A. E. and Aiken, L. S., 2013. Correcting injunctive norm misperceptions motivates behavior change: a randomized controlled sun protection intervention. *Health Psychology*, 32 (5), 551.
- Rimal, R. N. and Real, K., 2003. Understanding the influence of perceived norms on behaviors. *Communication Theory*, 13 (2), 184-203.
- Sandstrom, M., Makover, H. and Bartini, M., 2013. Social context of bullying: Do misperceptions of group norms influence children's responses to witnessed episodes? *Social Influence*, 8 (2-3), 196-215.
- Schreier, S. S., Heinrichs, N., Alden, L., Rapee, R. M., Hofmann, S. G., Chen, J., Oh, K. J. and Bögels, S., 2010. Social anxiety and social norms in individualistic and collectivistic countries. *Depression and anxiety*, 27 (12), 1128-1134.
- Sleeper, M., Balebako, R., Das, S., McConahy, A. L., Wiese, J. and Cranor, L. F., 2013. The post that wasn't: exploring self-censorship on facebook (pp. 793-802).
- Statista, Q., 2016. Number of monthly active Facebook users worldwide as of 4th quarter 2015.
- Steen-Johnsen, K. and Enjolras, B., 2016. The Fear of Offending: Social Norms and Freedom of Expression, 352.
- Sudman, S., Bradburn, N. M. and Schwarz, N., 1996. *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.
- Swire, B. and Ecker, U., 2018. ELEVEN Misinformation and Its Correction: Cognitive Mechanisms and Recommendations for Mass Communication. *Misinformation and mass audiences*. University of Texas Press, 195-211.
- Tanaka, Y., Sakamoto, Y. and Matsuka, T., 2013. Toward a Social-Technological System that Inactivates False Rumors through the Critical Thinking of Crowds (pp. 649-658): IEEE.
- Tandoc, E. C., Lim, D. and Ling, R., 2020. Diffusion of disinformation: How social media users respond to fake news and why. *Journalism*, 21 (3), 381-398.
- Tjosvold, D. and Sun, H. F., 2002. Understanding conflict avoidance: Relationship, motivations, actions, and consequences. *International Journal of Conflict Management*.
- Triandis, H. C., 1994. Culture and social behavior.
- Tully, M., Bode, L. and Vraga, E. K., 2020. Mobilizing Users: Does Exposure to Misinformation and Its Correction Affect Users' Responses to a Health Misinformation Post? *Social Media + Society*, 6.
- Valenzuela, S., Kim, Y. and Gil de Zúñiga, H., 2012. Social networks that matter: Exploring the role of political discussion for online political participation. *International Journal of Public Opinion Research*, 24 (2), 163-184.
- Van Houtte, M. and Stevens, P. A. J., 2008. Sense of futility: The missing link between track position and self-reported school misconduct. *Youth & Society*, 40 (2), 245-264.
- Vicol, D. O., 2020. Who is most likely to believe and to share misinformation? Full Fact.

- Vosoughi, S., Roy, D. and Aral, S., 2018. The spread of true and false news online. *science*, 359 (6380), 1146-1151.
- Vraga, E. K. and Bode, L., 2017. Using expert sources to correct health misinformation in social media. *Science Communication*, 39 (5), 621-645.
- Walter, N. and Murphy, S. T., 2018. How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs*, 85:3, 423-441.
- Wickström, G. and Bendix, T., 2000. The "Hawthorne effect"—what did the original Hawthorne studies actually show? *Scandinavian journal of work, environment & health*, 363-367.
- Zhang, Z.-X. and Wei, X., 2017. Superficial harmony and conflict avoidance resulting from negative anticipation in the workplace. *Management and Organization Review*, 13 (4), 795-820.
- Zhang, Z.-X., Zhang, Y. and Wang, M., 2011. Harmony, illusory relationship costs, and conflict resolution in Chinese contexts. In: Leung, A. K. y., Chiu, C.-y. and Hong, Y.-y., eds. *Cultural processes: A social psychological perspective*. New York, NY: Cambridge University Press, 188-209.