The Effects of Risk and Role on Users' Anticipated Emotions in Safety-Critical Systems

Yusuf Albayram[1(✉)], Mohammad Maifi Hasan Khan[1], Theodore Jensen[1], Ross Buck[2], and Emil Coman[3]

[1] Department of Computer Science and Engineering, University of Connecticut, Storrs, USA
{yusuf.albayram,maifi.khan,theodore.jensen}@uconn.edu

[2] Department of Communication, University of Connecticut, Storrs, USA
ross.buck@uconn.edu

[3] Health Disparities Institute, University of Connecticut Health Center, Hartford, USA
coman@uchc.edu

Abstract. Users of safety-critical systems often need to make risky decisions in real-time. However, current system designs do not sufficiently take users' emotions into account. This lack of consideration may negatively influence a user's decision-making and undermine the effectiveness of such a "human-computer collaboration." In a two-way, 2 (role: operator/system administrator) $\times$ 3 (risk level: high/medium/low) factorial study, we investigated the intensity of 44 emotions anticipated by 296 Mechanical Turk users who imagined being the (1) operator or (2) administrator of a drone system identifying (a) enemies on a battlefield, (b) illegal immigrants or (c) whale pods. Results indicated that risk level had a significant main effect on ratings of negative individualistic and negative prosocial emotions. Participants assigned to the high risk scenario anticipated more intense negative individualistic (e.g., nervous) and negative prosocial (e.g., resentful, lonely) emotions and less intense positive (e.g., happy, proud) emotions than participants assigned to the medium and low risk scenarios. We discuss the implications of our findings for the design of safety-critical systems.

Keywords: Emotions · Human-computer interaction
Decision-making · Risk

1 Introduction

Drone systems are increasingly being used for various purposes such as border patrol, battlefield monitoring, target tracking, and recreational activities. These systems can malfunction due to environmental factors, communication errors, or

D. Harris (Ed.): EPCE 2018, LNAI 10906, pp. 369–388, 2018.
https://doi.org/10.1007/978-3-319-91122-9_31

hardware and software failures, all of which may cause users to experience strong negative emotions (e.g., anger, anxiety, frustration, regret). Although there is a growing body of research showing that emotions strongly influence decision-making under risk and uncertainty [1–3], current safety-critical system designs do not consider users' emotions. This is likely to undermine effective decision-making, as strong emotions (e.g., regret, suspicion) can alter users' cognitive process [4].

While the role of emotion was long thought to be disruptive and contrary to models of decision-making, it is now understood that considering only the rational and cognitive is incomplete [5,6]. For instance, prior work in communication theory and psychology suggests that risky situations involve complex strong emotions (e.g., fear, suspicion, excitement) and that, if forewarned about what emotions to expect (i.e., emotional education), people are less surprised by their emotions [1,4,7]. This can allow for mindful processing of risks (e.g., emotional inoculation) [6]. Because of the risks faced by safety-critical system users, we argue that "emotional inoculation" is widely applicable to safety-critical human-computer interaction, and should be explicitly considered while designing user interfaces. A system that communicates about emotions can improve decision-making by allowing users to process the strong negative and positive emotions that arise in their safety-critical tasks. Before designing such systems, it is important to identify the relevant emotions.

As a first step towards this goal, we investigated the effect of risk level and role on users' anticipated emotions in a two-way, 2 (role: operator/system administrator) $\times$ 3 (risk level: high/medium/low) factorial experiment. We recruited 296 participants on Amazon's Mechanical Turk platform and provided them with a written description of one of six hypothetical scenarios where they were asked to imagine themselves as a drone operator or system administrator in a high, medium, or low risk scenario. Participants rated the anticipated intensity of 44 emotions in their scenario. Our findings show that risk level had a significant main effect on negative individualistic emotions and negative prosocial emotions. Participants in the high risk scenario expected more negative individualistic (e.g., nervous), more negative prosocial (e.g., resentful, lonely) and fewer positive (e.g., happy) emotions than participants in the medium and low risk scenarios. Insights gained in this study can enhance our understanding of the emotional aspects of decision-making in safety-critical human-computer interaction. The details of our study are presented in the following sections.

2 Background

2.1 Emotions in Decision-Making

Decision-making is the process of selecting a preferred option or course of action among a number of choices [8]. For a considerable time, decision-making was regarded by researchers as a predominantly cognitive process. According to utility theory, decision-makers evaluate the potential consequences of their options and choose the one they believe will yield the most beneficial result (i.e., the "utility-maximizing" alternative) [9]. Research on decision-making in the last couple of decades has shown that this view is incomplete. There is now a significant amount of psychological research demonstrating that emotions influence decision-making in various ways [4,7].

In a review of these works, Loewenstein and Lerner [4] note two different ways in which emotions enter into a decision: (1) *expected emotions* and (2) *immediate emotions*. *Expected emotions* are those that a decision-maker thinks they will experience as a consequence of some decision. Considered alongside the utility model, the decision-maker will evaluate the consequences of their options and choose that which they expect to maximize positive emotions and minimize negative emotions. *Immediate emotions* are those experienced at the time of decision-making.

Prior work suggests that *immediate emotions* and *expected emotions* are interconnected: *immediate emotions* can impact expectations about future emotions, while *expected emotions* that are anticipated by a decision-maker can influence their current emotional state [4]. For instance, studies have shown that if a decision-maker is presently experiencing positive emotions, his or her evaluation of certain options is likely to be more positive, while those experiencing negative emotions are likely to make more negative evaluations [10,11]. This is exemplified by a "hot/cold empathy gap," in which individuals in a "hot" emotional state (e.g., angry) have been observed to poorly predict their feelings or behavior when in a "cold" state (e.g., not angry) [12]. Additionally, findings that positive emotions broaden attentional focus while negative emotions narrow it [13,14] suggest that the valence and nature of an individual's *immediate emotions* influence their cognitive processing. These dynamics have clear implications for decision-making.

In situations involving risk and uncertainty, not only is there a potential increase in cognitive workload, but the effects of the decision-maker's emotions become more pronounced [1,3]. The "*risk as feelings hypothesis*" explores this notion to explain behavioral responses that differ from what individuals cognitively view as the best course of action. While moderately intense emotions tend to play an "advisory role," and their influence on an individual's judgment can often be limited [4,15], strong emotions generally exert more control over behavior. The "*risk as feelings hypothesis*" lends this to the role of "anticipatory" emotions such as fear, worry, and anxiety as inputs in the decision-making process. Specifically, there are a different set of determinants for cognitive evaluations of risk and emotional reactions to risks. While the former is influenced by factors such as outcome probability and severity, emotions are influenced more so by the vividness of imagined consequences or experience with certain outcomes. For instance, feelings about risk have been found to be insensitive to changes in probability, contrary to cognitive evaluations of risk [1].

Use of safety-critical systems is a high-risk, decision-making context where both moderate, advisory emotions and stronger emotions are likely to be at play.

2.2 Emotions in Human-Computer Interaction

Safety-critical system users such as drone operators and air traffic controllers often need to make decisions under uncertainty and time pressure. As wrong decisions may lead to serious consequences for people, property and the environment [16], users of such systems are likely to experience strong anticipatory emotions. Likewise, although the probability of the computer system failing is likely to be low, the potential negative consequences can be emotionally salient. Therefore, it is important to understand what specific emotions may be experienced by users.

Interaction with computers is often portrayed as a purely cognitive endeavor, given that the machines literally operate based on logic. However, recent research highlights the importance of emotional considerations in human computer interaction, wherein a computer that can recognize human emotion can appropriately respond its user's emotions, thus improving the user experience and outcomes of the interaction [17–20]. In one application, Jones and Jonsson [21] proposed an emotionally responsive car system that tracks the emotional state of a driver based on their speech. This information is then used to modify the car's navigational voice, which can relax a tense driver or make them happier about the current conditions. This can improve the driver's concentration and improve safety. This study reports promising results on the potential for emotions to be actively and effectively leveraged in safety-critical human-computer interaction.

Recently, Buck et al. [22] presented the User Affective eXperience (UAX) scale, measuring self-reported emotions that were anticipated in response to pop-up software update messages. They reported 4 latent factors (positive affect, anxiety, hostility and loneliness) which were found to be significantly different between a pressured condition (imagining working on an urgent and stressful task) and a relaxed condition (imagining surfing on the Web while relaxing). Their findings suggest that considering only emotional valence is inadequate, while distinguishing between individualist and pro-social emotions can paint a more thorough picture of the dynamics of affect-influenced decision-making in HCI.

It is fairly obvious that the stress associated with risky, safety-critical system use may cause a user to experience individualistic emotions such as anger or confusion. It is less clear for prosocial emotions, such as guilt and shame, which are those associated with adherence to social norms and group cooperation [23]. First, these are relevant in the drone context because of the presence of other people: system use can have direct consequences for people on the ground, while human operators and administrators work together on tasks with the system. Yet further, a substantial amount of research showing that humans respond socially to computer interaction partners [24,25] suggests that prosocial emotions may arise in the "group cooperation" between human members of the team and the computer system itself. Whereas Freedy et al. [26] sought to define better performance metrics for the unique "interaction of two cognitive systems" (i.e., the human and the computer), we argue that human emotions play an equally important role in the dynamics of such a "collaborative mixed initiative system".

For example, user emotions may contribute to their "trust" in an automated system, which has been found to influence reliance decisions [27]. Problems of automation *disuse*, in which operators do not use a system when it may help, and *misuse*, in which operators use a system when it is insufficient for some task, are well cited and have been linked to poor "calibration" of trust by the user ("undertrust" and "overtrust," respectively) [28]. Thus, several researchers have investigated the factors that influence a trust in automation, often varying system reliability and measuring trust with self-reports [29]. While it has been noted that there may be affective components of trust in addition to analogical ones, the role of emotions in trust decisions has not been sufficiently studied. Given that the consequences to poor trust calibration may be particularly severe with safety-critical systems, we argue that affective trust is highly influential on users' decision-making.

While some research efforts have investigated the influence of emotions in human-computer interaction (HCI), to the best of our knowledge, we are the first to investigate the effects of risk and role on users' anticipated emotions in the context of safety-critical drone applications. Specifically, this study expands upon Buck et al.'s work [22] and explores the anticipated intensity of 44 discrete emotions across various roles and risk levels with respect to a safety-critical drone system.

3 Methodology

3.1 Study Design

This study investigates how a safety-critical system users' anticipated emotions vary depending on their role and the criticality of the situation. Toward that, we designed six hypothetical scenarios involving drone operations. Among multiple possible safety critical technologies (e.g., smart grid, self-driving car, assisted robots, drones), this study uses drone because they are utilized for diverse applications (e.g., purely entertainment, border patrol, war).

The experiment was a 2 (role: operator/system administrator) $\times$ 3 (risk level: high/medium/low), between-subject factorial design where participants were randomly assigned to one of six hypothetical scenarios. Participants were asked to rate the anticipated intensity of 44 emotions while imagining themselves in their "risk level" and "role."

The two "roles" used in the study are as follows:

– ***System Administrator*:** The task involves managing a drone that is used by someone else (e.g., operator), and making sure the system is working/operating properly.
– ***System Operator*:** The task involves making decisions with and operating a drone that is overseen by system administrators.

The three "risk levels" used in the study are as follows:

- ***High Risk*: ** The drone was over a battlefield, and the decisions involve identifying enemy targets who may be innocent civilians.
- ***Medium Risk*: ** The drone was over a border region, and the decisions involve arresting suspected illegal immigrants who may be innocent citizens.
- ***Low Risk*: ** The drone was over the ocean, and the decisions involve identifying whale pods or non-interesting seals for a company.

The written descriptions of the scenarios were identical with the exception of the roles and risk level they mentioned, and are outlined in the Appendix. In particular, the hypothetical drone system had some operational instabilities that could cause negative performance. This information was intended to stimulate participants' emotional responses as they imagined making decisions in a safety-critical situation (i.e., with potentially dangerous consequences) with this imperfect system.

3.2 Survey

We designed a survey consisting of multiple parts as follows.

First, participants were asked to answer demographic questions (e.g., age, gender, and level of education) and report their level of computer proficiency. They were then shown a video about drones and their various applications. Following the video, participants were asked if they understood what drones are, and whether they had prior experience with drones (for either fun or professional reasons).

Subsequently, participants were randomly assigned to one of the six scenarios and, as an attention check, were asked to provide a written explanation of how the drone system is operated, how reliable it is, what their role and task was in the given scenario, and the risks associated with decisions they would have to make.

Finally, participants were asked to rate the expected intensity of 44 different emotions on a scale ranging from 1 (the least amount of intensity) to 7 (the greatest amount of intensity). The emotions were presented in the format "*I would feel* [***Emotion***]" and shown to participants in random order to avoid biasing them. These emotions were chosen to cover the broad range of emotional responses one could have while using a computer system [22,30,31]. The list of the 44 emotions can be seen in Table 3 in the Appendix.

We expected participants in the high risk scenario (i.e., identifying enemies on a battlefield) to report higher levels of negative emotions (e.g., nervous, anxious) than those in the medium risk (i.e., identifying illegal immigrants) and low risk (i.e., identifying whale pods) scenarios. Additionally, we expected the intensity of negative and positive emotions to vary between operator and administrator roles in the same scenario due to different responsibilities.

Moreover, prior work has found distinction between individualistic and prosocial emotions in response to pop-up software update warning messages [22]. In our hypothetical context, the distinction between individualistic and prosocial emotions may also be salient, given that (1) system failure could lead to negative consequences for other people and (2) the task involves collaboration with other people and the computer system itself. Thus, we expected to find differences in individualistic and prosocial emotions across risk levels and roles.

3.3 Participants

We recruited participants from Amazon's Mechanical Turk (MTurk) platform. We restricted participants to those 18 or older, currently living in the United States, having greater than 1000 approved HIT's (Human Intelligence Tasks), and having a HIT approval rate greater than 95%.

A total of 300 participants were recruited. We removed the responses of 4 participants who failed to properly answer the attention check question. Thus, a total of 296 valid responses were included in our analysis. Table 1 shows the distributions of participants among the six groups.

Table 1. 6 hypothetical scenarios: 2 roles (i.e., administrator and operator) and 3 risk levels (i.e., high, medium and low risk). The number of participants in each group is also shown.

	Number of participants	Role	Risk level
Scenario-1	48	System admin	High risk
Scenario-2	51	System operator	
Scenario-3	49	System admin	Medium risk
Scenario-4	49	System operator	
Scenario-5	50	System admin	Low risk
Scenario-6	49	System operator	

Participants took an average of 17.7 min ($Median$=14.8, SD=11.6 min) to complete the survey and were compensated with $3. The study was approved by the University's Institutional Review Board (IRB).

3.4 Demographics

Out of 296 participants who completed the survey, 158 (53.4%) were male. Participants' age ranged from 19 to 67 with an average of 33.5 years (median = 32, std = 9.4). All but 3 participants reported English as their native language.

In terms of education level, 89.8% of participants reported having some form of postsecondary education (e.g., college or university) while the most frequent reported education level was a 4-year college degree 43.2% (128). The breakdown of the other reported education levels is as follows: high school/GED (10.1%; 30), some college (23%; 68), 2 year college (14.9%; 44), master's degree (6.4%; 19), and doctoral or professional degree (2.4%; 7).

In terms of reported knowledge about computers in general, 9 (3.0%) participants identified themselves as "beginner," 5 (1.7%) as "novice," 90 (30.4%) as "competent," 150 (50.7%) as "proficient," and 42 (14.2%) as "expert." Moreover, 7 (2.4%) participants reported that they did not know what drones were before watching the video, while only one participant reported not knowing after watching the video. Overall, 39 (13.2%) participants reported having had experience with drones for either fun or professional reasons.

To examine demographic differences among the six groups, we performed an exploratory analysis with gender, age, level of education, knowledge about computers, and prior experience with drones. The results of the analysis revealed no significant differences in gender ($\chi^2(5) = 5.79$, $p = 0.32$), age ($\chi^2(5) = 4.93$, $p = 0.42$), education ($\chi^2(5) = 6.28$, $p = 0.27$), reported computer expertise ($\chi^2(5) = 7.86$, $p = 0.16$) or prior experience with drones ($\chi^2(5) = 5.12$, $p = 0.40$) across the six groups.

Based on our analysis, we concluded that the groups recruited were similar in terms of demographics.

4 Findings

We first performed an exploratory Principal Component Analysis (PCA) on the ratings of the 44 anticipated emotions. This analysis allowed us to cluster the emotions into groups (i.e., factors) and determine the characteristics of each. Subsequently, for each factor extracted, we performed a 2-way, 2×3 (role $\times$ risk level) Analysis of Variance (ANOVA). The details are presented below.

4.1 Factor Analysis

To assess the appropriateness of the collected emotion data for factor analysis, we first conducted several diagnostic tests using well-known sampling adequacy measures. Bartlett's test of sphericity measure is ($\chi^2(946) = 8725.2$, $p < 0.0001$) and the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy is 0.934. According to the Kaiser criterion, 0.9 and above reveals marvelous value [32], suggesting that our data was correlated and that the variability can be explained by common factors.

Subsequently, we conducted an exploratory PCA on the ratings of the 44 emotions and extracted 6 emotion factors based on the Kaiser criterion (i.e., K1 rule: retain factors if eigenvalue is greater than 1). However, as the Kaiser criterion often leads to substantial overfactoring [33], we also performed parallel analysis and determined the optimal coordinates. Briefly, parallel analysis calculates eigenvalues based on the same sample size and number of variables using sets of random data. Then, each *ith* eigenvalue obtained from the random data is compared with the *ith* eigenvalue produced by the actual data. Based on this comparison, the eigenvalue is retained if the eigenvalue expected from random data is greater than the eigenvalue calculated by the factor analysis. The optimal coordinate method uses linear regression to determine the coordinates where an eigenvalue diverges [34]. These two methods (i.e., parallel analysis and optimal coordinates) are widely used for determining the appropriate number of factors.

As shown in Fig. 1, both parallel analysis and optimal coordinates suggest extracting three factors for our data. Based on the aforementioned methods, we extracted three factors. These three factors predicted a cumulative total of 55.78% of the variance where factors 1, 2 and 3 explain 29.51%, 18.97%, and 7.29% of the variance, respectively.

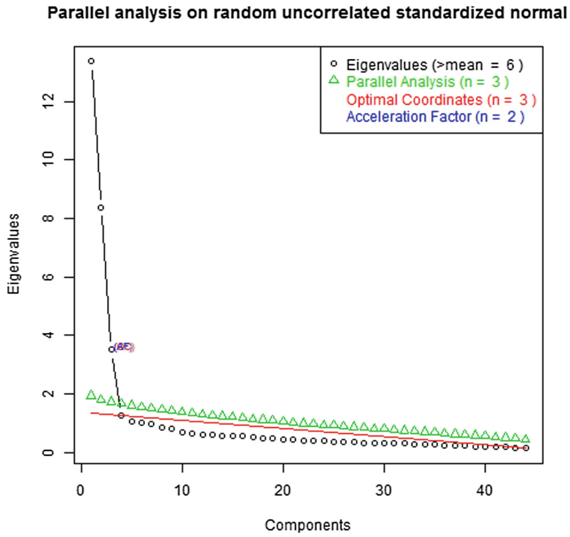

Fig. 1. Scree plot showing eigenvalues from the factor analysis, parallel analysis, optimal coordinates, and acceleration factor.

We used Varimax (orthogonal) as the rotation method, wherein prior work has considered items with a loading above 0.4 to be loaded on a factor [35]. Table 2 shows the rotated factor loadings of 44 emotions as well as the emotions belonging to each factor. Nineteen emotions such as angry, nervous and dismayed were included in Factor-1, which was labeled as "Negative individualistic" emotions. Fifteen emotions such as happy, welcomed and grateful were included in Factor-2, which was labeled as "Positive" emotions. Lastly, ten emotions such as scornful, disdainful and resentful were included in Factor-3, which was labeled as "Negative prosocial" emotions. These factors support those found in Buck et al.'s work [22] in the context of software update pop-up warnings, with our "Negative individualistic" corresponding to their "Anxious," our "Positive" to that of the same label, and "Negative prosocial" to the pair of factors "Lonely" and "Hostile."

For our three extracted factors, we also computed reliability measures using Cronbach's α. As shown in the second to last row of Table 2, all Cronbach's α values are higher than 0.7. According to McKinley et al. [36], α >0.6 indicates

Table 2. Factor loadings of the 44 emotions from the factor analysis. The highest factor loadings of each factor are highlighted in bold to facilitate visualization. The reliability measures (Cronbach's α and average inter-item correlation (IIC)) are also shown in the last two rows.

	Factor-1	Factor-2	Factor-3
Angry	**.775**		
Nervous	**.773**		
Dismayed	**.771**		
Anxious	**.768**		
Distraught	**.751**		
Ashamed	**.746**		
Down	**.732**		
Embarrassed	**.731**		
Afraid	**.731**		
Guilty	**.716**		
Sad	**.697**		
Freaked out	**.664**		
Depressed	**.649**		
Disgusted	**.632**		
Confused	**.592**		
Dazed	**.587**		
Hostile	**.528**	.475	
Isolated	**.514**	.484	
Surprised	**.444**		
Happy		**.776**	
Welcomed		**.769**	
Grateful		**.762**	
Admiring		**.760**	
Proud		**.758**	
Triumphant		**.758**	
Powerful		**.756**	
Secure		**.746**	
Trusting		**.744**	
Friendly		**.737**	
Cared-for		**.730**	
Respectful		**.717**	
Confident		**.681**	
Vigorous		**.672**	
Energetic		**.668**	
Scornful			**.789**
Disdainful			**.745**
Resentful			**.729**
Dishonored			**.716**
Contemptuous			**.709**
Humiliated			**.669**
Arrogant			**.651**
Lonely			**.602**
Insulted	.493		**.571**
Abandoned	.504		**.515**
Cronbach's alpha (α)	.946	.940	.886
IIC	.479	.511	.525

satisfactory internal reliability for all sub-scales. Finally, we calculated average inter-item correlation (IIC) values. As shown in the last row of Table 2, all the sub-scales are above 0.30, indicating "exemplary" reliability [37]. Based on our analysis, we concluded that each of the extracted factors had high reliability.

4.2 ANOVA Analysis

As we wanted to better understand how users might feel while using the drone system in different scenarios and roles, we performed a two-way, (2 × 3) ANOVA for each emotion factor extracted from the factor analysis. More specifically, the dependent variables for our three ANOVAs were negative individualistic emotions (factor-1), positive emotions (factor-2), and negative prosocial emotions (factor-3). We included risk level (high, medium, and low risk), role (system operator and system administrator), and their interaction effects as independent variables in each analysis. The details are presented below.

The ANOVA revealed that risk level had a significant main effect on negative individualistic emotions $F(2{,}290) = 6.8$, $p = .001$ and negative prosocial emotions $F(2{,}290) = 4.1$, $p = .017$. Participants assigned to the high risk scenario anticipated stronger negative individualistic emotions (e.g., nervous, confused) and negative prosocial emotions (e.g., resentful, lonely), but weaker positive (e.g., happy, grateful) emotions than those assigned to the medium risk and low risk scenarios. More specifically, participants in the high risk scenario ($Mean = 3.54$, $SD = 1.27$) rated higher negative individualistic emotions than participants in the medium scenario ($Mean = 3.11$, $SD = 1.32$) and the low risk scenario ($Mean = 2.88$, $SD = 1.26$). A series of post-hoc pairwise comparisons using Bonferroni correction revealed that the difference in ratings between the

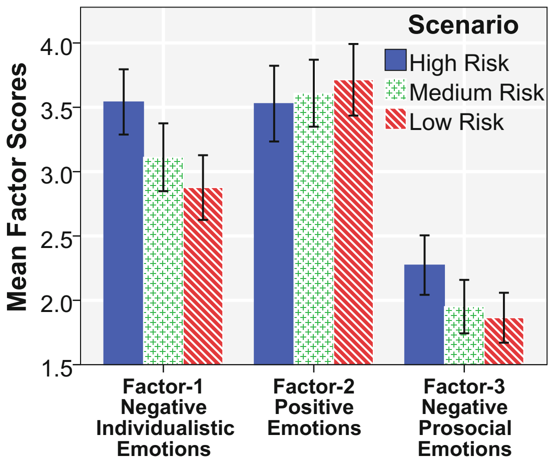

Fig. 2. Mean factor scores for the three risk levels (high risk/medium risk/low risk) for each factor. 95% confidence intervals are also included.

high risk and the low risk scenarios was significant ($p < .001$). Similarly, participants in the high risk scenario ($Mean = 2.27$, $SD = 1.16$) rated negative prosocial emotions higher than participants in the medium risk scenario ($Mean = 1.95$, $SD = 1.04$) and the low risk scenario ($Mean = 1.86$, $SD = 0.97$). A series of post-hoc pairwise comparisons using Bonferroni correction revealed that the difference in ratings between the high risk and low risk scenarios was significant ($p < .021$). Although those in the high risk scenario ($Mean = 3.53$, $SD = 1.48$) rated lower levels of positive emotions than participants in the medium risk scenario ($Mean = 3.61$, $SD = 1.30$) and low risk scenario ($Mean = 3.71$, $SD = 1.40$), the difference in ratings among the three risk levels was not statistically significant. The mean factor scores for the three risk levels are shown in Fig. 2.

The ANOVA also revealed that there was no significant main effect on emotions due to role. The mean factor scores for the two roles (operator/system administrator) can be seen in Fig. 3.

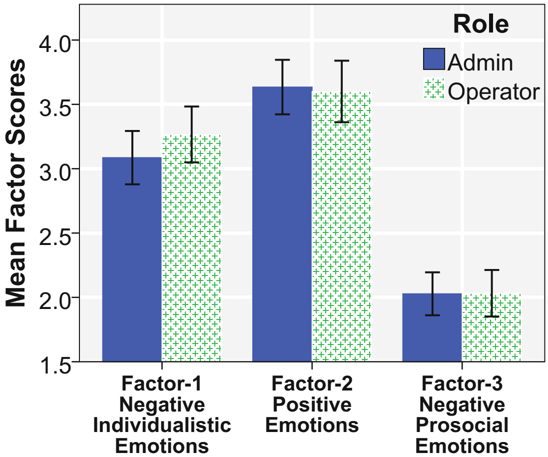

Fig. 3. Mean factor scores for the two roles (system administrator/system operator) for each factor. 95% confidence intervals are also included.

5 Discussion

Despite a growing body of literature demonstrating the significant role of emotions in the decision-making process, we have a relatively limited understanding of the specific emotions relevant to high risk decision-making. As safety-critical technologies such as drones and self-driving cars become more prevalent, so will the high-risk decisions to which their users must attend. To gain insight into the effect of risk level and role on safety-critical system users' emotions, we asked participants to imagine themselves as a drone operator or system administrator in a high, medium, or low risk scenario. They then rated the expected intensity of 44 emotions while imagining the scenario.

We found that participants in the high risk scenario reported higher levels of negative individualistic emotions (e.g., angry, nervous), negative prosocial emotions (e.g., scornful, resentful) and lower levels of positive emotions (e.g., happy, grateful) than participants assigned to the medium and low risk scenarios. These differences were significant between high and low risk participants for both negative prosocial and negative individualistic emotions. These findings suggest that, unsurprisingly, use of safety-critical systems may involve strong negative emotions. The notion that computers are cognitive entities, with which interaction should be non-emotional in order to be efficient and successful, may be particularly destructive in this context. A lack of acknowledgment by the system may not only alter a user's decision-making, but lead to stronger negative emotions that impact later interaction.

Developing emotion-aware communication strategies by detecting users' emotions during system operations can reduce the potentially harmful effects of negative emotions. Specifically, teaching users to recognize their emotions (emotional education) may enable them to act more mindfully, and help to lessen the potential negative effects of strong emotions on decision-making (emotional inoculation) [6]. We argue that "emotional inoculation" is particularly applicable in the safety-critical domain, such as our hypothetical drone system. Communicating with users about emotions they may experience while using a system can positively contribute to both their decision-making outcomes and their perceptions of the system. Future work should test the effectiveness of safety-critical system interfaces that incorporate emotional inoculation via different types of messages and in various decision-making contexts. Furthermore, "emotional inoculation" and "emotional education" can be incorporated into training materials for safety-critical system users (e.g., drone operators). Using virtual simulators in realistic scenarios, such training systems could inform operators about the emotions they might experience during certain points of system use (e.g., feeling nervous and anxious during a time-sensitive task) and the nature of the specific emotions in such situations (e.g., prosocial vs. individualistic, or positive vs. negative). This can help prepare operators to regulate their reactions under time pressure and stress while performing complex safety-critical tasks [38].

These kinds of emotional communication can help to improve a user's trust calibration. Prior work has found that happiness, as well as "liking" a system influence reliance [39]. These affective aspects may help to explain changes in trust over the course of a human-computer interaction [29,39,40]. Future work should explore how the negative individualistic and negative prosocial emotions associated with safety-critical system use factor into trust evaluations and reliance decisions, as well as how an understanding of these emotions can be leveraged to improve system design and trust calibration.

We also found that, at the same risk level, the intensity of emotion factors differed (see Figs. 2 and 3). Negative prosocial emotions had the lowest mean intensity in all risk levels and roles, whereas positive emotions and negative individualistic emotions generally had higher intensity. Though prosocial emotions were not felt as strongly by participants, we observed that their anticipated intensity differed between high and low risk level participants. It appears that users are not just thinking about themselves with their use of the drone system, but about the involvement of others. This result is in line with research demonstrating the relevance of both individualistic and prosocial emotions in the context of pop-up security messages [22]. In the drone context, prosocial emotions could have been associated with (1) people on the ground who may have been impacted by the drone, (2) other human collaborators, or (3) the computer system itself. The latter is supported by research demonstrating social responses to computers by human users [24]. Future work could shed light on the specific effect that the computer itself has on user emotions by investigating how factors of the system and its interface influence the intensity of prosocial emotions, relative to differences in the context of system use.

Lastly, we found that for all the three factors, the interaction between risk level and role was not significant. This indicates that participants' emotions were more likely to be influenced by the criticality of the situation rather than their assigned role. It is possible that participants in operator and administrator roles in the same scenario considered the level of risk the same, and thus the role to which they were assigned did not make a strong contribution to their overall feelings. Such a difference may be more pronounced in a lab setting where participants interact directly with a system. If the user's role on a task-oriented team is more linked to their actions, then emotions may be impacted by their level of responsibility for team success.

5.1 Limitations

While this study provides insights about the effects of risk and role on users' emotions, there are several limitations in this work.

First, we used hypothetical (i.e., artificial) scenarios in which participants rated how they would expect to feel as the operator or administrator of a drone system. Given the lack of actual interaction with a computer system, it may have been difficult for participants to anticipate the emotions they would experience. Moreover, this could contribute to misinterpretations of the degree of risk. For example, some participants in the low risk condition (i.e., identifying whale pods) may have considered the situation to be very risky, since failure could have caused "job loss." Nevertheless, even in this artificial scenario-based methodology, our results revealed considerably diverse ratings of emotions depending on the group to which participants were assigned.

Second, we recruited participants from the MTurk platform. Although MTurk allows for recruiting larger and more diverse populations in terms of age, education level and ethnicity compared to samples from specific subpopulations (e.g., students enrolled in a psychology class) [41,42], it is hard to verify the attentiveness of MTurk users. To filter out responses that demonstrated a lack of understanding of the scenario, we included an attention check question in the survey.

Lastly, since our study was survey-based, emotional states of participants were measured via self-reports. Though our data provides insight into the role of "anticipated emotions" in a risky human-computer interaction, it needs further validation given that individuals may have difficulty predicting their emotional states [43]. To develop a more thorough understanding of user's emotions, future studies should investigate the somatic components (e.g., facial expressions and the heartbeat) [44] of "immediate emotions" in studies involving actual human interaction with a computer system.

We believe that this work is a useful starting point for research on the role of emotions in decision-making with safety-critical systems, which has important implications for system interface design. We encourage future work to investigate the specific factors that influence user emotions (e.g., risk and the nature of consequences, organizational structure, system features) as well as the influence that different types of emotions have on decision-making, behavior, and system performance.

6 Conclusion

This study aimed to understand the role of emotions in decisions at various risk levels and responsibilities with respect to a safety-critical system. Participants were asked to rate the intensity with which they would feel 44 emotions while imagining using a drone system in one of six hypothetical scenarios where they were asked to imagine themselves as a drone operator or system administrator in a high, medium, or low risk scenario. We found that participants assigned to the high risk scenario anticipated more intense negative individualistic, negative prosocial and less intense positive emotions than participants assigned to medium and low risk scenarios. We strongly believe that insights gained in this work will enable researchers to develop more effective emotionally-aware communication strategies for safety-critical systems.

Acknowledgments. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-15-1-0490.

Appendix

The Descriptions of the Scenarios

Through a two-way 2 (role: system operator/system administrator) × 3 (risk level: high risk/medium risk/low risk) factorial design experiment, participants were assigned to one of the six hypothetical scenarios. Depending on the assigned role and risk level, participants were shown one of the phrasings separated by vertical bars (|) below. For instance, participants assigned to the system operator role in the high risk scenario were shown *($Role_{opr}$) and ($Risk_{high}$)*, while participants assigned to the system administrator role in the low risk scenario were shown *($Role_{adm}$) and ($Risk_{low}$)*, and so on. The entire written description of scenarios are outlined below.

Now, imagine that [$Role_{opr}$: *there are system administrators who are*] | [$Role_{adm}$: *you are the system administrator who is*] responsible for:

- Making sure that the software of the system that is used to operate the drone remotely is up-to-date.
- Making sure that the hardware of the system is up-to-date.
- Troubleshooting of the system if the performance is not acceptable.
- Performing preventative maintenance of the system.

However, despite [$Role_{opr}$: *their*] | [$Role_{adm}$: *your*] best effort, the system is not perfectly reliable and the system occasionally experiences the followings due to software bugs/hardware failures:

- The system occasionally crashes due to some unknown reasons and takes 2 min to reboot, making the system unavailable, and the timing/frequency of the crash is unpredictable.
- The system occasionally becomes very slow (e.g., freezes for 10 s at a time) due to unknown software/hardware bugs.
- The system occasionally drops video frames due to communication errors.
- Different hardware components of the system rarely fails (e.g., once every 6 months).

Now, imagine that [$Role_{opr}$: *you are asked to use the system to make*] | [$Role_{adm}$: *the drone that you are responsible for managing is going to be used by someone else (e.g., operator) whose*] decisions involve identifying:

- $Risk_{high}$: enemy targets in a battlefield where there may be also innocent civilians.
- $Risk_{med}$: arresting or not arresting suspected illegal immigrants who may be innocent citizens in a border region.
- $Risk_{low}$: whale pods or non-interesting seals in the ocean for a company.

Table 3. Participants were asked to rate the expected intensity of these 44 emotions on a scale ranging from 1 (the least amount of intensity) to 7 (the greatest amount of intensity).

Emotions
1. I would feel TRUSTING (e.g. because the system has given an opportunity to respond)
2. I would feel HAPPY (e.g., because I am informed of actual system states)
3. I would feel CONFIDENT (e.g., because I am informed of actual system states)
4. I would feel SECURE (e.g., because I am informed of actual system states)
5. I would feel SAD (e.g., because the system is not performing as expected)
6. I would feel DEPRESSED (e.g., because the system is not performing as expected)
7. I would feel DOWN (e.g., because the system is not performing as expected)
8. I would feel AFRAID (e.g., because the system is not performing as expected)
9. I would feel NERVOUS (e.g., because the system is not performing as expected)
10. I would feel ANXIOUS (e.g., because the system is not performing as expected)
11. I would feel ANGRY (e.g., because the system is not performing as expected)
12. I would feel INSULTED (e.g., because the system is not performing as expected)
13. I would feel HOSTILE (e.g., because the system is not performing as expected)
14. I would feel SURPRISED (e.g., because one does not expect the interruption)
15. I would feel DAZED (e.g., because one does not expect the interruption)
16. I would feel CONFUSED (e.g., because one does not expect the interruption)
17. I would feel FREAKED OUT (e.g., because one does not expect the interruption)
18. I would feel DISGUSTED (e.g., because the system is not performing as expected)
19. I would feel DISMAYED (e.g., because the system is not performing as expected)
20. I would feel DISTRAUGHT (e.g., because the system is not performing as expected)
21. I would feel CARED-FOR (e.g., because I am informed of actual system states)
22. I would feel FRIENDLY (e.g., because I am informed of actual system states)
23. I would feel WELCOMED (e.g., because I am informed of actual system states)
24. I would feel POWERFUL (e.g., because I am warned and can respond)
25. I would feel ENERGETIC (e.g., because I am warned and can respond)
26. I would feel VIGOROUS (e.g., because I am warned and can respond)
27. I would feel ISOLATED (e.g., because my response may be inadequate)
28. I would feel LONELY (e.g., because my response may be inadequate)
29. I would feel ABANDONED (e.g., because my response may be inadequate)
30. I would feel PROUD (e.g., because I am warned and can respond)
31. I would feel TRIUMPHANT (e.g., because I am warned and can respond)
32. I would feel ARROGANT (e.g., because I am warned and can respond)
33. I would feel ASHAMED (e.g., because my response may be inadequate)
34. I would feel GUILTY (e.g., because my response may be inadequate)
35. I would feel EMBARRASSED (e.g., because my response may be inadequate)
36. I would feel SCORNFUL (e.g., because the system state is fine)
37. I would feel CONTEMPTUOUS (e.g., because the system state is fine)
38. I would feel DISDAINFUL (e.g., because the system state is fine)
39. I would feel HUMILIATED (e.g., because the system state is fine)
40. I would feel DISHONORED (e.g., because the system state is fine)
41. I would feel RESENTFUL (e.g., because the system state is fine)
42. I would feel GRATEFUL (e.g., because the system has given an opportunity to respond)
43. I would feel RESPECTFUL (e.g., because the system has given an opportunity to respond)
44. I would feel ADMIRING (e.g., because the system has given an opportunity to respond)

References

1. Loewenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N.: Risk as feelings. Psychol. Bull. **127**(2), 267 (2001)
2. Schlösser, T., Dunning, D., Fetchenhauer, D.: What a feeling: the role of immediate and anticipated emotions in risky decisions. J. Behav. Decis. Mak. **26**(1), 13–30 (2013)
3. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econom.: J. Econom. Soc. **47**, 263–291 (1979)
4. Loewenstein, G., Lerner, J.S.: The role of affect in decision making. Handb. Affect. Sci. **619**(642), 3 (2003)
5. Buck, R., Davis, W.A.: Marketing risk: emotional appeals can promote the mindless acceptance of risk. In: Roeser, S. (ed.) Emotions and risky technologies, pp. 61–80. Springer, Dordrecht (2010). https://doi.org/10.1007/978-90-481-8647-1_4
6. Buck, R., Ferrer, R.: Emotion, warnings, and the ethics of risk communication. In: Roeser, S., Hillerbrand, R., Sandin, P., Peterson, M. (eds.) Handbook of Risk Theory, pp. 693–723. Springer, Dordrecht (2012). https://doi.org/10.1007/978-94-007-1433-5_27
7. Lerner, J.S., Li, Y., Valdesolo, P., Kassam, K.S.: Emotion and decision making. Ann. Rev. Psychol. **66**, 799–823 (2015)
8. Wilson, R.A., Keil, F.C.: The MIT Encyclopedia of the Cognitive Sciences. MIT press, Cambridge (2001)
9. Harless, D.W., Camerer, C.F.: The predictive utility of generalized expected utility theories. Econom.: J. Econom. Soc. **62**, 1251–1289 (1994)
10. Clore, G.L.: Cognitive phenomenology: feelings and the construction of judgment. Const. Soc. Judgm. **10**, 133–163 (1992)
11. Clore, G.L., Schwarz, N., Conway, M.: Affective causes and consequences of social information processing. Handb. Soc. Cogn. **1**, 323–417 (1994)
12. Loewenstein, G.: Out of control: visceral influences on behavior. Organ. Behav. Hum. Decis. Process. **65**(3), 272–292 (1996)
13. Basso, M.R., Schefft, B.K., Ris, M.D., Dember, W.N.: Mood and global-local visual processing. J. Int. Neuropsychol. Soc. **2**(3), 249–255 (1996)
14. Conway, M., Giannopoulos, C.: Dysphoria and decision making: limited information use for evaluations of multiattribute targets. J. Pers. Soc. Psychol. **64**(4), 613 (1993)
15. Forgas, J.P.: Mood and judgment: the affect infusion model (AIM). Psychol. Bull. **117**(1), 39 (1995)
16. Knight, J.C.: Safety critical systems: challenges and directions. In: Proceedings of the 24th International Conference on Software Engineering, ICSE 2002, pp. 547–550. ACM, New York (2002)
17. Brave, S., Nass, C.: The Human-Computer Interaction Handbook, pp. 81–96. L. Erlbaum Associates Inc., Hillsdale (2003)
18. Brave, S., Nass, C., Hutchinson, K.: Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. Int. J. Hum.-Comput. Stud. **62**(2), 161–178 (2005)
19. Picard, R.W., Klein, J.: Computers that recognise and respond to user emotion: theoretical and practical implications. Interact. Comput. **14**(2), 141–169 (2002)
20. Klein, J., Moon, Y., Picard, R.W.: This computer responds to user frustration: theory, design, and results. Interact. Comput. **14**(2), 119–140 (2002)

21. Jones, C., Jonsson, I.-M.: Using paralinguistic cues in speech to recognise emotions in older car drivers. In: Peter, C., Beale, R. (eds.) Affect and Emotion in Human-Computer Interaction. LNCS, vol. 4868, pp. 229–240. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85099-1_20
22. Buck, R., Khan, M., Fagan, M., Coman, E.: The user affective experience scale: a measure of emotions anticipated in response to pop-up computer warnings. Int. J. Hum.-Comput. Interact. **34**, 1–10 (2017)
23. Bowles, S., Gintis, H.: Prosocial Emotions. The Economy As an Evolving Complex System III, pp. 339–366. Santa Fe Institute, Santa Fe (2005)
24. Reeves, B., Nass, C.I.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, Cambridge (1996)
25. Nass, C., Moon, Y.: Machines and mindlessness: social responses to computers. J. Soc. Issues **56**(1), 81–103 (2000)
26. Freedy, A., DeVisser, E., Weltman, G., Coeyman, N.: Measurement of trust in human-robot collaboration. In: 2007 International Symposium on Collaborative Technologies and Systems, CTS 2007, pp. 106–114. IEEE (2007)
27. Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., Beck, H.P.: The role of trust in automation reliance. Int. J. Hum.-Comput. Stud. **58**(6), 697–718 (2003)
28. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse, abuse. Hum. Factors **39**(2), 230–253 (1997)
29. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**(1), 50–80 (2004)
30. Buck, R., Anderson, E., Chaudhuri, A., Ray, I.: Emotion and reason in persuasion: applying the ari model and the casc scale. J. Bus. Res. **57**(6), 647–656 (2004)
31. Kay, R.H., Loverock, S.: Assessing emotions related to learning new software: the computer emotion scale. Comput. Hum. Behav. **24**(4), 1605–1623 (2008)
32. Kaiser, H.F.: An index of factorial simplicity. Psychometrika **39**(1), 31–36 (1974)
33. Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., Strahan, E.J.: Evaluating the use of exploratory factor analysis in psychological research. Psychol. Methods **4**(3), 272 (1999)
34. Raîche, G., Walls, T.A., Magis, D., Riopel, M., Blais, J.G.: Non-graphical solutions for Cattell's scree test. Methodol.: Eur. J. Res. Methods Behav. Soc. Sci. **9**(1), 23–29 (2013)
35. Stevens, J.P.: Applied Multivariate Statistics for the Social Sciences. Routledge, London (2012)
36. McKinley, R.K., Manku-Scott, T., Hastings, A.M., French, D.P., Baker, R.: Reliability and validity of a new measure of patient satisfaction with out of hours primary medical care in the united kingdom: development of a patient questionnaire. BMJ: Br. Med. J. **314**(7075), 193–198 (1997)
37. Robinson, J.P., Shaver, P.R., Wrightsman, L.S.: Criteria for scale selection and evaluation. Meas. Personal. Soc. Psychol. Attitudes **1**(3), 1–16 (1991)
38. Luini, L.P., Marucci, F.S.: Prediction-confirmation hypothesis and affective deflection model to account for split-second decisions and decision-making under pressure of proficient decision-makers. Cogn. Technol. Work **17**(3), 329–344 (2015)
39. Merritt, S.M.: Affective processes in human-automation interactions. Hum. Factors **53**(4), 356–370 (2011)
40. Merritt, S.M., Ilgen, D.R.: Not all trust is created equal: dispositional and history-based trust in human-automation interactions. Hum. Factors **50**(2), 194–210 (2008)

41. Chandler, J.J., Paolacci, G.: Lie for a Dime: when most prescreening responses are honest but most study participants are impostors. Soc. Psychol. Personal. Sci. **8**(5), 500–508 (2017)
42. Landers, R.N., Behrend, T.S.: An inconvenient truth: arbitrary distinctions between organizational, mechanical turk, and other convenience samples. Ind. Organ. Psychol. **8**(2), 142–164 (2015)
43. Picard, R.W., Picard, R.: Affective Computing, vol. 252. MIT press, Cambridge (1997)
44. Han, S., Lerner, J.S., Sander, D., Scherer, K.: Decision making. In: Sander, D., Scherer, K.R. (eds.) The Oxford Companion to Emotion and the Affective Sciences. Oxford University Press, Oxford (2009)