Factors Affecting Compliance with Alerts in the Context of Healthcare-related Emergencies

Manasvi Kumar

Department of Management Information Systems, Eller College of Management, University of Arizona, USA

manasvik@email.arizona.edu

Gondy Leroy

Department of Management Information Systems, Eller College of Management, University of Arizona, USA

gondyleroy@arizona.edu

Abstract:

This study is a conceptual replication of the study by Han et al. (2015) in which the authors evaluated factors affecting students' compliance with emergency instructions during campus emergencies. The current study focused on broader public health-related emergencies using eight scenarios evaluated by Amazon Mechanical Turk (AMT) participants. Analysis on the aggregated data showed that subjective norm and trust in information quality positively affected intention to comply with instructions, consistent with the original study. Three follow up analyses provide more nuanced results. First, an abridged dataset was created by filtering out participants whose reasons for non-immediate compliance weren't related to verifying information and then complying. Analysis on this dataset showed that subjective norm no longer positively affected intention to comply. Second, scenarios used in the study were grouped by characteristics such as development speed, frequency, and area affected, and the analysis was redone. Factors affecting intention to comply immediately changed based on the characteristic, with subjective norm positively affected intention to comply in slow-developing scenarios, scenarios that affect at a limited area, and commonly occurring scenarios, while trust in information quality affects the other scenarios. Third, recall of information from the notification was collected from participants and analyzed. Results show that participants who chose to comply immediately recalled more information than others. Our replication study shows some support for the original conclusions; however, the broader setting and more nuanced analyses show also differences between both studies.

Keywords: Surveys, Analytical modeling, Computer-mediated communication

1　Introduction

Mass notification systems for campus-related emergencies grew in significance after the 2007 Virginia Tech shooting. These notification systems are seen as an interplay between two factors (Gow et al., 2009) – the technical factor (notification capabilities) and the social factor (communication considerations), and improving their effectiveness is important. Studies focusing on the social factor aim to understand system effectiveness from the perspective of the university public involved (Madden, 2015), while there are numerous studies that focus on improving the technology involved in notification communication. Campus notifications are studied under the umbrella of the extreme events domain (Han, 2015).

This study is a conceptual replication of the study by Han et al. (2015): the hypotheses are retained but the context, treatments and evaluation are broader and applicable to a general public (Dennis & Valacich, 2014). The original study aimed to identify factors affecting compliance with campus-related emergency notifications by developing a model and hypotheses adopted from Etzioni's compliance theory (Etzioni 1961a, Etzioni 1961b, Etzioni 1961c). The model and hypotheses were designed to accommodate for emergency notifications sent by a normative organization. The current study also focuses on studying factors affecting compliance in the case of emergencies, but in the context of public health. Since both universities (campus-related emergencies) and governments (public health-related emergencies) are organizations which are normative in nature, the original hypotheses remain applicable. The original study used a swine-flu scenario for its "health-related incidents." We present several different public health-related scenarios and aim to generalize results from the original study by analyzing our data at an aggregate level. We present additional insights by categorizing scenarios based on three properties: development speed, area affected and frequency. These results could help generate insights for new scenarios based on their relevance to a category. As public health-related emergencies can affect a more generalized population compared to campus-related emergencies, participants of the current study were recruited from Amazon Mechanical Turk (AMT). A comparison between the two studies is summarized in Table 1.

Table 1. Comparison between the two studies		
Category	**Current Study**	**Han et al., 2015**
Research Question	What are the factors that influence the intention to comply in a healthcare related emergency?	What are the factors that influence the intention to comply in a campus emergency?
Research Method	Survey (modified to include additional questions and free recall of information)	Survey + Focus group
Scenarios	Healthcare-related emergencies	Campus emergencies
Analysis	Aggregate level and Category level	Scenario level
Participants	Amazon Mechanical Turk workers	Students
Gender	Male:134(49%) Female:138 (51%)	Male:308(38%) Female:513 (62%)
Education level	Undergraduate and below: 230 (85%) Graduate and above: 42 (15%)	Undergraduate: 670 (82%) Graduate and above: 151 (18%)

The model used in both studies builds on Etzioni's compliance theory and previous studies on trust (Choudhury & Karahanna, 2008; Dashti et al., 2010). Etzioni viewed compliance as a relationship between power employed by superiors in an organization to control their subordinates and the orientation of these subordinates towards this power. Organizations are distinguished based on the types of power used to control employees. The original study adapted Etzioni's compliance theory to an emergency notification system context characterized by employment of normative power (no physical or financial punishment for noncompliance). Normative power was adapted as subjective norm (SN), which referred to the perceived social and peer pressure while performing certain behaviors. Other types of power, such as coercive power and remunerative power, were adapted as perceived safety threat (PST) and perceived financial threat (PFT). Based on previous studies on trust and past experience (Bonsall & Parry, 1991), information quality trust (IQT) and past experience (PE) were also considered for the model with the latter being a control

variable. The model adopted in the original study and results obtained for the public health-related incidents are shown below (Figure 1). The current study adopts the same model and hypotheses:

Hypothesis 1 (H1): Subjective norm positively affects the intention to comply immediately.

Hypothesis 2 (H2): Perceived safety threat positively affects the intention to comply immediately.

Hypothesis 3 (H3): Perceived financial threat positively affects the intention to comply immediately.

Hypothesis 4 (H4): Information quality trust positively affects the intention to comply immediately.
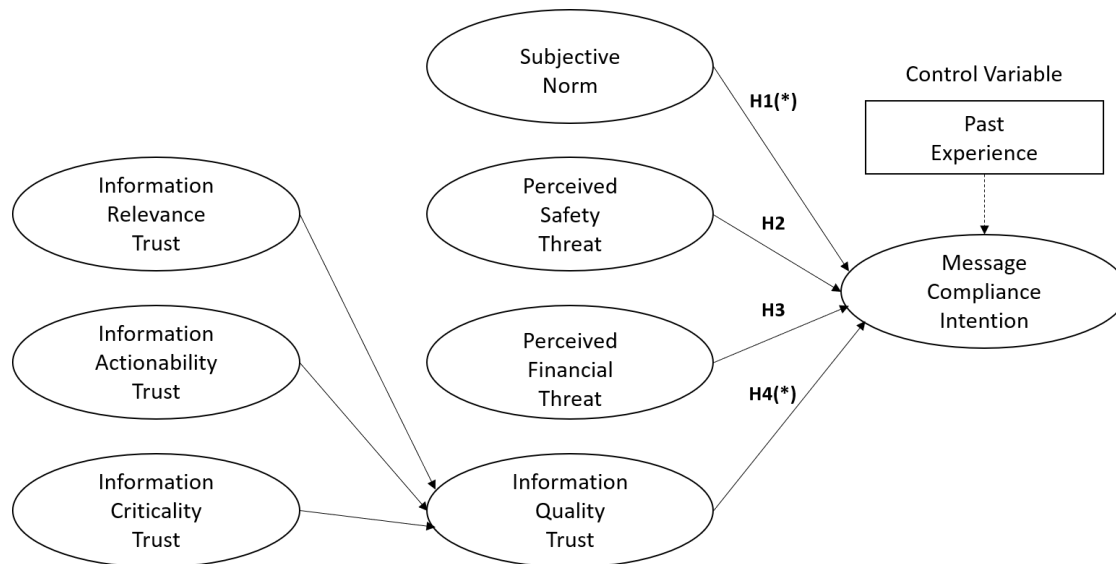

Figure 1. Research model from the original study with results for public health-related scenarios (* denotes that the hypothesis was supported) (Adapted from Han et. al, 2015)

2   Research Methodology

As in the original study, a survey was conducted to identify factors affecting compliance in various scenarios. However, no focus groups were conducted for qualitative analysis. Instead, all participants of the current study were asked to explain their reason for non-compliance and to recall the notification information at the end of the survey. In addition, the study focused on a broader topic (public health emergencies) with participants selected from a broader group (adult respondents in the US).

2.1   Instrument Development

The questions asked were adapted from the original study with changes made when necessary.

2.1.1   Scenario Design

The study utilized a scenario-based survey to measure variables that affected intention to comply. Eight public health scenarios were chosen based on characteristics specified in the disaster matrix (Table 2). The study extended the framework from the original study (developed by Zdziarski et al., 2007). The scenarios were categorized by a public health expert. All emergencies were categorized by their development speed (fast or slow), their frequency of occurrence (common or rare), and the extent of the geographical area affected (limited or wide).

The explanations and scenarios presented to each survey participant are listed in Appendix A. The explanations were intended to educate participants about the threats associated with a scenario. For example, a rail accident scenario usually affects the passengers on board the train and the local site of impact. If the train carried poisonous gases, however, the threat would be different and the effect more widespread. Such threats are detailed in the explanation so that all participants have the same interpretation

of the scenario. At the end of the explanation, a notification was provided to the participant with recommended actions. For example, for rail accidents, the notification read, "Train crash 25 miles north from downtown releases harmful gases. If you could be at risk, report to local authorities soon." All notification texts designed for the current study were maximally 128 characters as in the original study. A minor distinction between the two is the use of a less demanding action in the current study (i.e., report "soon" as opposed to report "immediately"). This was changed to provide consistency across scenarios and allowing that, realistically, immediate compliance is not always possible.

Table 1. Disaster matrix showing scenarios selected based on characteristics				
Development Speed	**Slow-Developing Emergencies**		**Fast-Developing Emergencies**	
Frequency	**Rare**	**Common**	**Rare**	**Common**
Limited Area Affected	Rail accident	Salmonella	Wildfire	Tornado
Wide Area Affected	Measles	Influenza	Earthquake	Hurricane

2.1.2   Survey

The participants were presented with a consent form at the start of the survey followed by demographic questions. Then, participants completed one or more scenarios. Each scenario and explanation were followed by 24 questions to measure the five independent variables in the original study. These questions were adapted from the original study but slightly modified to suit the new context. For example, referent groups such as "Colleagues" and "Community" replaced "University Officials" and "Professors" to accommodate for a general population. At the end, the participants were asked to indicate their intent to comply with the notification if they were the person facing the scenario. For example, with the earthquake scenario, the question was "If you were Gary, what would you do?" The answer options were "Comply immediately," "Verify first and then comply," or "Ignore." An additional question helped identify participants who could not relate to the scenario: "Can you imagine a similar emergency happening to you/someone in the US?" Responses from participants who did not relate to the scenario were discarded (similar to the original study).

In addition to the questions adapted from the original survey, three new questions were added. The first question asked participants who chose to comply after verification to justify their rationale (a text box was provided for participants to explain their rationale). The second question was added to determine whether the participant paid adequate attention while taking the survey ("Which emergency was this about?"). Attention-check questions such as this are commonly used with Amazon Mechanical Turk. Finally, the third question asked participants to recall all information from the notification. The rationale behind this question was to analyze the similarities of the responses to the actual notification for various cases (different categories, age groups, compliance behaviors, etc.). (See Appendix B for the questions asked in the survey and the summary statistics for various responses obtained.)

2.2   Instrument Validation

A reliability and validity analysis of the constructs was performed first. As in the original study, survey data were aggregated and SmartPLS 2.0 (Ringle et al., 2005) used to perform this analysis. Demographic data were used to check for homogeneity within the samples using chi-square tests based on gender ratios and academic status (undergraduate and below or graduate and above). For scenarios with a different rate of development (slow- or fast-developing), no significant differences were observed in gender ratio ($\chi^2 = 0.49$, $p = 0.48$) or academic status ($\chi^2 = 0.79$, $p = 0.37$). Similarly, there were no significant differences when categorizing based on area affected (wide or limited) for gender ratio ($\chi^2 = 0.21$, $p = 0.65$) and academic status ($\chi^2 = 2.19\text{e-}06$, $p = 0.99$). No significant differences were observed when categorizing based on frequency of scenarios for gender ratio ($\chi^2 = 0.45$, $p = 0.50$) and academic status ($\chi^2 = 1.41$, $p = 0.23$). Thus, the samples chosen were deemed homogenous on the basis of gender and academic status.

2.2.1   Validation of Reflective Constructs

For each reflective construct, average variance extracted (AVE) values and individual item loadings were calculated to check for reliability and convergent validity. Table 3 shows that item loadings for the reflective

constructs were greater than 0.6. This score is considered sufficiently high to establish reliability (Backhaus et al., 2003). Table 4 shows that the Cronbach's alpha values were greater than 0.7 and the AVE values were greater than 0.5, which places them both above their recommended values (Fornell & Larcker, 1981) to establish the convergent validity of constructs. To establish discriminant validity, factor analysis values (Table 3) were compared between constructs to ensure that the difference was at least 0.10, as recommended (Gefen and Straub, 2005). The square root of the AVE values for each construct were compared to the correlations in the correlation matrix. Table 4 shows that correlations on the diagonal were greater than the off-diagonal elements, thereby establishing discriminant validity.

Table 3. Factor loadings (reflective constructs)		
	PST	**PFT**
PST1	0.992	0.568
PST2	0.940	0.519
PFT1	0.470	0.942
PFT2	0.591	0.889

Table 4. Latent variable correlations, AVE, Composite reliability, and Cronbach's alpha (reflective constructs)					
	PST	**PFT**	**AVE**	**Composite Reliability**	**Cronbach's Alpha**
PST	0.967		0.839	0.912	0.942
PFT	0.567	0.916	0.934	0.966	0.812

2.2.2   Validation of Formative Constructs

For formative constructs, convergent validity and reliability testing were not conducted because they are not relevant here. To test discriminant validity, the significance levels of item weightings were examined and confirmed to be greater than 0.05, as suggested by Diamantopoulos and Winklhofer (2001). Variance inflation statistics revealed no multicollinearity between items, since all values were lower than 3 (Diamantopoulos & Siguaw, 2006.)

3   Results

Data obtained from the survey were analyzed at the aggregate level to generalize for the public health-related incidents and at category level to determine the impact of characteristics of various scenarios on compliance intention. A mixed-effects model was used to identify the factors that influenced immediate compliance (denoted by 1 in the model) versus compliance after verification (denoted by 0). The model was run using R (Version 3.6.1), utilizing the "lme" function from the "nlme" package. Latent variable scores derived using SmartPLS provided input data for the independent variables of the model. The dependent and independent variables were the same as those in the original study. In the model, the independent variables of the original study were used as fixed effects. As some participants completed more than one survey (multiple scenarios), "Participant" was considered as a random effect to account for non-independence among responses. Similar to the original study, data collected through the survey was self-reported and was collected within a similar period of time. Hence, common method bias was a potential issue (Podsakoff et al., 2003). Several tests were conducted to ensure that common method bias did not the affect results of the study. Similar to the original study, procedural remedy was also used to reduce common method bias. These results are summarized in Appendix C.

3.1   Participants

This study focused on broader public health scenarios, not campus incidents, and the responders were recruited on AMT to provide representation. The participants were compensated $0.50 per scenario completed. They could complete multiple scenarios.

There were 394 surveys completed by 272 participants, with three participants completing all eight scenarios. Of these, 23 participants chose to "Ignore" the notification when asked about their intent to comply and these data were discarded from the analyses below. When asked if they could relate to the alert provided, 35 participants answered "No" and their answers were also discarded (as in the original study). Since the participants of the current study were AMT workers, we discarded an additional 13 surveys where participants failed the attention-check question (a common practice). The resulting 323 surveys were deemed usable. Table 5 shows the demographic data. Participants were mostly white, had a bachelor's degree, and were between 31 and 40 years old. There was no clear majority among the participants in terms of gender.

Table 5. Demographics of participants (N=272)	
Characteristics	**Frequency (Percentage)**
Gender	
Male	134 (49.26)
Female	138 (50.74)
Race	
American Indian Native Alaskan	1 (0.37)
Asian	20 (7.35)
Black or African American	19 (6.99)
Native Hawaiian or Other Pacific Islander	0 (0)
White	223 (81.99)
More than one race	9 (3.31)
Education	
Less than High School	1 (0.37)
High School Diploma	70 (25.74)
Associates degree	57 (20.96)
Bachelor's degree	102 (37.50)
Master's degree	33 (12.13)
Doctorate	9 (3.31)
Age	
Younger than 30 years old	79 (29.04)
31 to 40 years old	104 (38.24)
41 to 50 years old	46 (16.91)
51 to 60 years old	25 (9.19)
61 to 70 years old	17 (6.25)
71 or more	1 (0.37)

Of the 323 surveys, 174 (53.87%) responses indicated that participants chose to comply with the notification *after* verification. The rationale provided by the participants was categorized by the authors (see Table 6). Most of these participants (119) chose to assess the authenticity of the information through other sources such as social media, news, or friends. The authors of the original study defined intention to "verify first and

then comply" as the intent to contact other people or sources to verify the information before complying. Based on the answer analysis, the participants in the current study who provided a similar rationale were classified as "Strategy 1." Sample responses included the following: "I would check the internet for more information on the incident," "Turn on the TV or radio to see if it's on the news," and "I would Google it to see how severe the consequences might be." Some participants (11) chose to assess the situation without external help, while a third group (44) expressed diverse concerns and chose passive strategies. These participants were classified into "Strategy 2" and "Strategy 3" respectively.

Table 6. Strategies followed by participants who chose to comply after verification		
Strategy	**Example**	**Number (Percentage)**
Strategy 1: Assess authenticity of information through other sources	"I would check other media outlets to see what is going on…"	119 (68.39%)
Strategy 2: Assess the situation without any external help	"I like to use my own judgment to assess the situation, unless the warning is very specific…"	11 (6.33%)
Strategy 3: Other strategies	"Wait cautiously"	44 (25.28%)

3.2   Dataset Creation

To acknowledge these differences in actions related to "Verify then comply," two datasets were created (Table 7): an unabridged dataset and a verify-first dataset. The unabridged dataset included data from all 323 participants, which were deemed usable similar to the original study. The verify-first dataset included data from the 268 participants who chose to comply immediately or those who would verify using Strategy 1 (which represents most closely the definition of intention to "verify first and then comply" from the original study). Thus, analysis on this dataset would most closely represent the vision of the authors from the original study. The mixed-effects analysis was therefore conducted on each dataset separately (unabridged and verify-first), on both the aggregate and category level.

Table 7. Datasets created for analysis			
Dataset	**Comply Immediately (N)**	**Verify first then comply (N)**	**Total N**
Unabridged dataset	Participants who chose to comply immediately (149)	Participants who chose either Strategy 1, 2, or 3 (174)	323
Verify-first dataset	Participants who chose to comply immediately (149)	Participants who chose either Strategy 1(119)	268

Since some participants completed more than one survey, a mixed-effect model was used by including the independent variables from the original study as fixed effects and participant as a random effect to resolve the non-independence of some responses. The effect size for each of our models was determined using Cohen's f statistic. Twelve of the 14 models used in this study detected at least one parameter with a medium effect size (f>0.25 and 90% Confidence Interval), and two models detected parameters with small effect sizes (Maximum f=0.19 and 0.23 for Rare scenario categories of the Unabridged and Verify-first datasets). Table 8 shows the sample size for each scenario and category.

Prior to mixed-effects analysis, correlations among the variables were calculated at the aggregate level and for all categories. Table 9 shows that there were four pairs of variables (Ex: PST and IQT in the Fast-developing category) whose correlation exceeded the threshold value of 0.60, a threshold used in previous studies (Hair et al., 1995; Peng & So, 2002) that indicated possible multicollinearity. The variance inflation factor (VIF) values were calculated for these and found to be below three, which is commonly accepted as an indication that no multicollinearity problem exists (James et al., 2014). Thus, no variables were discarded and all five variables were used to perform the analysis at both scenario and category levels.

Table 8. Sample size for each scenario and category				
Development Speed	**Slow-Developing Emergencies (128)**		**Fast-Developing Emergencies (195)**	
Frequency	**Rare (174)**	**Common (149)**	**Rare**	**Common**
Limited Area Affected (151)	Rail accident (28)	Salmonella (28)	Wildfire (40)	Tornado (55)
Wide Area Affected (172)	Measles (33)	Influenza (39)	Earthquake (73)	Hurricane (27)
Sample size mentioned in brackets i.e., (n)				

Table 9. Correlation among the variables for different scenarios (unabridged)					
Scenario		**IQT**	**PFT**	**PST**	**SN**
All scenarios	IQT	1.000			
	PFT	0.259**	1.000		
	PST	0.591**	0.567**	1.000	
	SN	0.466**	0.187**	0.382**	1.000
Slow	IQT	1.000			
	PFT	0.218*	1.000		
	PST	0.581**	0.423**	1.000	
	SN	0.535**	0.172	0.357**	1.000
Fast	IQT	1.000			
	PFT	0.271**	1.000		
	PST	**0.618****	0.458**	1.000	
	SN	0.428**	0.199**	0.464**	1.000
Limited	IQT	1.000			
	PFT	0.289**	1.000		
	PST	**0.666****	0.546**	1.000	
	SN	0.527**	0.262**	0.573**	1.000
Wide	IQT	1.000			
	PFT	0.256**	1.000		
	PST	0.539**	0.532**	1.000	
	SN	0.425**	0.136	0.234**	1.000
Rare	IQT	1.000			
	PFT	0.284**	1.000		
	PST	**0.647****	0.532**	1.000	
	SN	0.467**	0.227**	0.349**	1.000
Common	IQT	1.000			
	PFT	0.286**	1.000		
	PST	0.497**	**0.637****	1.000	
	SN	0.466**	0.157	0.413**	1.000
(* indicates $p < 0.05$, ** indicates $p < 0.01$, bold font indicates values above 0.60 threshold)					

3.3   Aggregate Level Results

The mixed-effects model was run for the unabridged and the verify-first dataset. Table 10 shows the results, with "Positive" denoting a positive, significant effect between the variable and the intention to comply immediately, i.e., the hypothesis from the original study holds. N/S (not significant) denoted a lack of such an effect.

In the unabridged dataset, subjective norm and information quality trust were positively related to the intention to comply immediately. This paralleled the results for health-related incidents in the original study. In the verify-first dataset, only information quality trust was positively related with the intention to comply immediately.

Table 10. Summarized results of analyses for datasets at aggregate level			
	Original Study	**Current Study**	
	Health Related-Incident	**Unabridged dataset**	**Verify-first dataset**
Past Experience	N/S	N/S	N/S
SN (H1)	**Positive**	**Positive**	N/S
PST (H2)	N/S	N/S	N/S
PFT (H3)	N/S	N/S	N/S
IQT (H4)	**Positive**	**Positive**	**Positive**

3.4   Category Level Results

All scenarios were categorized based on three characteristics and analyzed separately (e.g., fast or slow development speed). This approach differed from the original study, where each scenario was analyzed individually. The summarized results obtained are presented in Table 11. Similar analysis for the verify-first dataset is presented in Appendix D. The results for the unabridged and verify-first dataset are the same for the development speed category. Results vary in the categories of area affected and frequency of occurrence when the two datasets were analyzed using the mixed effect model.

Table 11. Summarized results of analyses for categories of unabridged dataset						
	Development speed		**Area affected**		**Frequency of occurrence**	
	Slow	**Fast**	**Limited**	**Wide**	**Rare**	**Common**
Past Experience	N/S	N/S	N/S	N/S	N/S	N/S
SN (H1)	**Positive**	N/S	**Positive**	N/S	N/S	**Positive**
PST (H2)	N/S	N/S	N/S	N/S	N/S	N/S
PFT (H3)	N/S	N/S	N/S	N/S	N/S	N/S
IQT (H4)	N/S	**Positive**	N/S	**Positive**	**Positive**	N/S

When comparing scenarios with different rates of development, subjective norm positively affected intention to comply immediately for the slow-developing scenarios, while information quality trust did have an impact for the fast-developing scenarios. When comparing scenarios based on the area affected, a similar result was obtained. For scenarios which affected a limited area, subjective norm positively affected intention to comply, while information quality trust positively affected intention to comply when the area affected was wide. Comparing scenarios based on their frequency of occurrence also offered similar results, with subjective norm positively affecting the intention to comply in common scenarios and information quality trust positively affecting intention to comply in rarer scenarios.

Table 12 shows detailed results of the mixed-effects model. The effects of perceived safety threat, perceived financial threat, and past experience on intention to comply were not significant in any category. The correct

classification rate was greater than the baseline (constant classification model) in all categories, indicating that the mixed-effects models outperformed the baseline model while predicting intention to comply. Similar detailed analysis for the verify-first dataset is summarized in Appendix D.

Table 12. Detailed results of mixed-effects analysis for unabridged dataset and its categories								
			Development rate		Area affected		Frequency of occurrence	
Variables	Statistics	Overall	Slow	Fast	Limited	Wide	Rare	Common
Past Experience	β	0.067	0.196	-0.034	0.023	0.070	0.037	0.053
	SE	0.068	0.122	0.082	0.103	0.098	0.126	0.060
Subjective Norm	β	0.067*	0.119*	0.068	0.155**	0.010	0.030	0.135**
	SE	0.030	0.048	0.040	0.048	0.041	0.042	0.045
Perceived Safety Threat	β	0.033	0.078	-0.033	-0.023	0.020	0.033	-0.023
	SE	0.036	0.050	0.046	0.058	0.049	0.053	0.056
Perceived Financial Threat	β	-0.043	-0.084	0.025	-0.001	-0.071	-0.811	0.025
	SE	0.030	0.045	0.037	0.043	0.043	0.043	0.049
Information Quality Trust	β	0.118**	0.051	0.133	0.098	0.163**	0.151**	0.087
	SE	0.034	0.055	0.041**	0.051	0.046	0.049	0.046
Participant (Intercept)	Variance	0.11	0.1038	0.133	0.982	0.079	0.111	0.689
	Std. Dev.	0.332	0.322	0.365	0.313	0.283	0.333	0.262
-2LL		421.097	172.598	270.862	208.607	246.247	250.318	206.548
Response ratio (Comply immediately: Verify then comply)		149:174	51:77	98:97	72:79	77:95	88:86	61:88
Immediate Compliance Rate		46.13	39.84	50.26	47.68	44.77	50.57	40.94
(* indicates p<0.05, ** indicates p< 0.01)								

3.5　Mean Contrast Analysis

To understand how each of the referent groups influenced immediate compliance, a mean contrast analysis (similar to the analysis performed in the original study) was performed and the results are summarized in Table 13. Mean contrasts were computed by calculating the difference of mean values of motivation to comply for each group (obtained from questions in the survey) from responses of participants who chose immediate compliance.

The results (Table 13) are similar to the health-related incidents of the original study. For example, influence exerted by family was greatest among all groups in the current study, similar to "Parents" in the original study. Colleagues exerted the least influence among all groups in the current study (similar to "Professors" in the original study). While there was no consistent trend among the other referent groups, the results show that other people exerted more influence than friends, who in turn exerted more influence than the community.

Table 13. Mean contrast of motivations to comply based on referent groups						
Scenario		**Family**	**Friends**	**Colleagues**	**Community**	**Other people**
All scenarios	Family	-				
	Friends	0.38	-			
	Colleagues	0.95**	0.58*	-		
	Community	0.66**	0.29	-0.29	-	
	Other people	0.19	-0.19	-0.77**	-0.48	-
Slow	Family	-				
	Friends	0.14	-			
	Colleagues	0.33	0.20	-		
	Community	0.17	0.04	-0.16	-	
	Other people	0.02	-0.12	-0.31	-0.16	-
Fast	Family	-				
	Friends	0.5	-			
	Colleagues	1.27**	0.78**	-		
	Community	0.92*	0.42	-0.36	-	
	Other people	0.28	-0.22	-1**	-0.64*	-
Limited	Family	-				
	Friends	0.31	-			
	Colleagues	0.81**	0.32	-		
	Community	0.63	-0.15	-0.18	-	
	Other people	0.15	-0.18	-0.65*	-0.47	-
Wide	Family	-				
	Friends	0.44**	-			
	Colleagues	1.09	0.65	-		
	Community	0.70	0.26	-0.39	-	
	Other people	0.22	-0.22	-0.87*	-0.48	-
Rare	Family	-				
	Friends	0.5	-			
	Colleagues	0.97**	0.47	-		
	Community	0.60	0.10	-0.36	-	
	Other people	0.08	-0.42	-0.89**	-0.52	-
Common	Family	-				
	Friends	0.20	-			
	Colleagues	0.93**	0.74*	-		
	Community	0.75*	0.56	-0.18	-	
	Other people	0.34	0.15	-0.59	-0.41	-
(* indicates $p < 0.05$, ** indicates $p < 0.01$)						

3.6   Free Recall Analysis

An extension suggested in the original study was to investigate effective ways to communicate with the respondents. To establish a baseline on how well participants understood the notification provided in the current study, each participant was asked to recall information about the notification. Recall was calculated using an in-house developed program that provides exact and semantic recall measures. The program first

performed spelling corrections to ensure comparability among respondents (many AMT workers use spell check, but not all). Then responses were compared to the original scenario content and similarity measures calculated. Exact recall values captured overlap in exact words between the response and notification, while semantic recall values captured overlap in words that were semantically similar. The numbers are average per group for words (AWR), nouns (ANR) and verbs (AVR).

Table 14 summarizes the results. The survey data was divided into several groups based on the scenarios assigned as well as actual survey responses and demographics. For each group, the mean value for metrics (AWR, AVR, ANR) for exact recall and semantic recall are reported. Average cosine similarity (ACS) of a response was calculated by computing the average of cosine similarity of word embeddings in the participants' response compared to the original notification. For all these measures, higher values represent a greater match between the participants' response and notification provided. Significance levels calculated using a conservative Bonferroni correction are also reported. Since 42 comparisons were made in each set (Different Scenarios and Different Groups), an alpha of 0.001 is required for significance.

Table 14. Free recall results comparing participants' responses and original notifications (AWR – Average words recalled, ANR – Average nouns recalled, AVR – Average verbs recalled, ACS – Average cosine similarity)								
Characteristic	Groups	Exact recall			Semantic recall			
		AWR	ANR	AVR	AWR	ANR	AVR	ACS
Free Recall for Groups characterized by Scenarios								
Development speed	Slow	3.15	1.78	0.77	4.51	2.38	0.94	0.147
	Fast	3.26	1.78	0.93	4.01	2.16	1.09	0.151
Area affected	Limited	3.09	1.84	0.66*	4.05	2.36	0.83*	0.137**
	Wide	3.33	1.73	1.05*	4.34	2.14	1.21*	0.16**
Frequency of occurrence	Common	2.87	1.46*	1.17**	4	1.97	1.44**	0.148
	Rare	3.52	2.06*	0.6**	4.38	2.48	0.68**	0.15
Free Recall for Groups characterized by Survey response / Demographic data								
Intention to comply	Comply Immediately	3.3	1.83	0.89	4.32	2.33	1.04	0.151
	Verify then comply	3.15	1.75	0.85	4.11	2.17	1.02	0.148
Gender	Male	3.44	1.98	0.91	4.42	2.41	1.06	0.149
	Female	3.01	1.6	0.83	4.01	2.09	1.01	0.15
Academic status	Education below graduate	3.22	1.79	0.85	4.25	2.28	1.02	0.149
	Graduate and above	3.2	1.76	0.96	3.93	2.02	1.09	0.151
(* indicates p < 0.05, ** indicates p < 0.01 for t test between differences of means after Bonferroni correction)								

Table 14 can help answer whether the type of scenario or the intent of the participant or other demographic data impacted information recalled by the participant. If a group receives higher scores on all metrics compared to the other, it indicates greater information recall. Grouping participants by scenarios provided ambiguous results for information recall as no group outperforms the other on all metrics. Similar results are obtained when grouping participants by demographic data like gender or academic status. When grouping participants by response on the intention to comply, it is found that participants who chose to comply immediately recall more information across all metrics.

4 Discussion and Conclusion

The purpose of this study was to replicate the study by Han et al. (2015) in the context of public health using eight public health-related scenarios and a survey with AMT participants. While the original study investigated several scenarios separately, the current study was performed at two levels – an aggregate level which cumulatively analyzed all scenarios together under the umbrella of public health, and a category level which combined scenarios based on their characteristics. Furthermore, the participants were asked to mention their rationale for compliance after verification (if required) and the responses obtained were analyzed. While most participants (68.39%) insisted that they wanted to verify the information provided through other sources (Strategy 1), a few participants gave other reasons. Thus, two datasets were created to perform analysis, and each of these datasets were analyzed at both levels.

Results of the original study for health-related incidents differed from the aggregate level results of the current study. Both H1 (subjective norm) and H4 (information quality) were supported in the original study and for the unabridged dataset, while only H4 was supported for the verify-first dataset. A possible explanation is that this may be an effect of the dataset composition. In the unabridged dataset, individuals with and without high subjective norm are included and a portion of these emphasize information quality. When reducing the datasets to those whose chose to verify first, the data are restricted to those with high pressure to follow an official message and a portion who emphasize information quality. However, this observed difference between the two datasets deserves further study.

A category level analysis was performed in the current study as opposed to scenario level analysis. Categories were created based on the area affected, development speed, and frequency of occurrence of any scenario. While comparing scenarios based on development speed, H1 was supported for slow-developing scenarios, while H4 was supported for fast-developing scenarios. This indicates that when time is scarce, trusting information provided in the notification is key to immediate compliance. When given adequate time, trust was no longer an important factor since participants could easily verify information provided prior to compliance. In this case, complying with social and peer pressure to follow expected optimal behavior, i.e., subjective norm, impacted immediate compliance. Table 12 shows that 51 out of 128 (39.8%) of participants chose to comply immediately in the slow-developing scenarios, while 98 out of 195 (50.3%) did so in fast-developing scenarios. Thus, when faced with slow-developing scenarios, the participants chose to make well-informed decisions, and immediate compliance rates were lower than usual. Comparing scenarios based on frequency of occurrence of scenarios yielded intuitive results. H1 was supported for commonly occurring scenarios while H4 was supported for rare scenarios. When scenarios are rare, preexisting knowledge pertaining to the scenario may be less and hence trusting information provided in the notification determines immediate compliance. In commonly occurring scenarios where expected behavior is well known, the perceived social and peer pressure to conform, i.e., subjective norm, determines immediate compliance. Comparing scenarios based on area affected indicated that H4 was supported for widely affected areas while H1 was supported only for scenarios affecting a limited area. It is possible that for scenarios which affect wide areas, information quality trust determines immediate compliance as appropriateness of action suggested in the notification could differ from area to area. In scenarios which affect a limited area and localized strategies are suggested, pressure to conform is greater and hence subjective norm determines immediate compliance.

The immediate compliance rate at the aggregate level for the unabridged dataset was 46.13% (149 out of 323 participants as indicated in Table 12). This is similar to the result for Health-related incidents (47.56%) in the original study. Thus, minor changes in notification text between the two studies did not majorly affect immediate compliance rates.

5 Limitations and Future Research

The current study is a conceptual replication of the original study with notable differences and limitations. Only one method to collect data (survey) was utilized without conducting an additional focus group. The current study aimed to identify factors affecting immediate compliance with emergency notifications relating to public health in the US.

Results obtained from the current study indicate mixed support for results from the original study. Results from the unabridged dataset were consistent with the original study, while those from the verify-first dataset differed. Results obtained also differed among the two datasets when analyzed at the category level. Current

events related to the spread of misinformation or "fake news" may impact the thinking of people about government-authorized notifications in general.

A key result from the free recall analysis was that participants who chose to comply immediately obtained greater scores on all metrics for both exact and semantic recall as opposed to those who chose to verify then comply. While this trend was observed based on the average values of responses, the differences between these scores weren't statistically significant. Future studies could attempt to recreate the study using a larger participant pool to check for significant differences. If true, a possible research question following this could be to understand the direction of causal relationship e.g., Is someone more likely to comply immediately if they can recall instructions or vice versa. Understanding this relation could provide practical insights to those who design the emergency notifications. For example, it is possible that participants who understood the instructions clearly were able to recall them and chose to comply immediately without seeking further help. If that is the case, notifications could be designed with an extra emphasis on being easily understood by many people.

5.1   Theoretical Implications

As is the case with replications, the main theoretical contribution of this study involves testing the generalizability of the methodology of the framework used in the original study. Our results indicate mixed support for the original study based on datasets created from additional information collected in the study. It was observed that *subjective norm* was not an indicator of compliance when the survey instrument was adjusted for a more precise definition of compliance after verification. Similar differences were observed when the data were analyzed at the category level. It is possible that other constructs outside the scope of this study could affect intention to comply. For example, previous studies by the authors of the original paper utilized constructs such as media richness (Han et al., 2011) to explain factors influencing compliance.

5.2   Practical Implications

As the current study analyzes intent to comply in the context of public health-related incidents, there are relevant practical implications. The study helps differentiate between factors which affect compliance in different scenarios and could lead to better strategies or customized alerts. For example, we found that in scenarios which occur rarely, information quality trust is a positive indicator for intention to comply. This indicates that a notification designed for this scenario should focus on emphasizing relevance, criticality, and actionability. Thus, by categorizing scenarios, we provide practitioners with relevant indicators to focus on while devising a notification. This applies to new scenarios as well, i.e., a new outbreak, since data was analyzed using properties of scenarios as opposed to the scenarios themselves.

Our study also indicates that several participants who chose to verify and then comply may want to do something other than contact their friends, family etc. For example, many participants suggested that they would assess the situation on their own. While it may not be possible to provide individualized feedback using emergency notifications, newer technologies such as crisis warning applications (Fischer et al., 2019), interactive messaging (e.g., with artificial chat bots to handle large demand), or use of video and media to facilitate remote verification could be deployed to increase compliance rates.

Acknowledgments

References

Backhaus, K., Erichson, B., Plinke, W., & Weiber, R. (2003). *Multivariate Analysemethoden* (10th ed.). Berlin: Springer.

Bonsall, P., & Parry, T. (1991). Using an interactive route-choice simulator to investigate driver's compliance with route guidance device. *Transportation Research Record* (No. 1306), 59-68.

Choudhury, V., & Karahanna, E. (2008). The relative advantage of electronic channels: A multidimensional view. *MIS Quarterly* 32(1), 179-200.

Dashti, A., Benbasat, I., & Burton-Jones, A. (2010). Trust, felt trust, and e-government adoption: A theoretical perspective. Paper presented at the *JAIS Theory Development Workshop 10*(83).

Dennis, A., & Valacich, J. (2014). A Replication manifesto. *AIS Transactions on Replication Research 1*, 1–5.

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative versus reflective indicators in organizational measure development: A comparison and empirical illustration. *British Journal of Management 17*, 263-282.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research 38*(2), 269-277.

Etzioni, A. (1961a). An analytical classification: Coercive and utilitarian organizations. In Etzioni, A., *A Comparative Analysis of Complex Organizations.* New York: Free Press, 23-39.

Etzioni, A. (1961b). Compliance as a comparative base. In Etzioni, A., *A Comparative Analysis of Complex Organizations.* New York: Free Press, 3-22.

Etzioni, A. (1961c). Compliance, goals, and effectiveness. In Etzioni, A., *A Comparative Analysis of Complex Organizations.* New York: Free Press, 71-88.

Fischer, D., Hattori-Putzke, J. and Fischbach, K. (2019). Crisis Warning Apps: Investigating the Factors Influencing Usage and Compliance with Recommendations for Action. Paper presented at the *52nd Hawaii International Conference on System Sciences*, Hawaii, HI.

Fornell, C., & D. F. Larcker (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* 18(39), 39–50.

G. A. Gow, T. McGee, D. Townsend, P. Anderson and S. Varnhagen. (2009). Communication technology, emergency alerts, and campus safety. *IEEE Technology and Society Magazine*, vol. 28, no. 2, pp. 34-41.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate Data Analysis with Readings* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Han, W., Ada, S., Sharman, R., Rao, H. R., & Brennan, J. (2011). Critical success factors to improve compliance with campus emergency notifications. Paper presented at the *17th Americas Conference on Information Systems,* Detroit, MI.

Han, W., Ada, S., Sharman, R., & Rao, H. R. (2015). Campus emergency notification systems: An examination of factors affecting compliance with alerts. *MIS Quarterly 39(4)*, 909-929.

Indushobha C. S., Saggi, N., and Pindaro, D. (2010). An empirical analysis of the business value of open source infrastructure technologies. *Journal of the Association for Information Systems,* 11(11), 708-729.

Lindell, M. K., and Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology,* 86(1), 114-121.

Nov, O., & Ye, C. (2008). Users' personality and perceived ease of use of digital libraries: The case for resistance to change. *Journal of the American Society for Information Science and Technology,* 59(5), 845-851.

Peng, C. J., & So, T. S. H. (2002). Logistic regression and reporting: A primer. *Understanding Statistics 1*(1), 31-70.

Podsakoff, P., MacKenzie, S., & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology 88*(5), 879-903.

Ringle, C. M., Wende, S., & Will, A. (2005). *SmartPLS 2.0.* Hamburg, Germany: University of Hamburg.

Zdziarski, E., Dunkel, Z., & Rollo, J. (2007). *Campus Crisis Management: A Comprehensive Guide to Planning, Prevention, Response, and Recovery*. San Francisco: Jossey-Bass.

Appendix A: Scenarios and their explanations.

Table A1. Scenarios presented to the participants of the study	
Scenario	**Explanation**
Earthquake	Surviving an earthquake can be very challenging. Even if one manages to survive the disaster without sustaining any injuries, there are other challenges in the aftermath of the event. Psychological and emotional stress factors aside, there are problems such as scarcity of food and water resources, which could be equally dangerous. An earthquake has rocked the city. Gary, one of the survivors living in the city, receives the following message from the notification system: "Earthquake affects several parts of the city. As water supply is compromised, residents are advised to drink bottled water only." Gary realizes that he does not have packaged drinking water with him.
Wildfire	Wildfires can cause significant damage to or totally destroy homes and could result in direct loss of life to humans and animals. Apart from this, they also release harmful gases into the air which lower air quality and cause massive damage to a person's lungs and eyes. Will came back home from his tiring road trip across the country. He receives a message from the emergency system: "Wildfire reported 50 miles from the city. If you have stinging eyes or issues with breathing normally, contact local authorities soon." Will realizes that he was driving through the area about an hour ago and has discomfort in his eyes.
Measles	Measles is an airborne infectious disease which spreads by the contagious measles virus. Despite its contagious nature, vaccination has proven to successfully prevent the large-scale spread of the disease. However, young children and unvaccinated adults are still at risk as recent outbreaks have shown. Laura was waiting for a bus to get back home from work. She received a message from the emergency system: "Measles outbreak reported in the city. If you are coughing, have a sore throat or skin rashes, see a healthcare provider soon." Laura has observed mild skin rashes developing on her skin in the last few days and is unsure whether she was vaccinated.
Rail accident	Rail accidents are fairly common in the United States and the damage is usually locally contained. However, in certain cases, the Graniteville train crash in 2005, the train transports toxic gases which are released into the air as a result of the crash. Dealing with these incidents requires swift action from the authorities and people alike. Karen came back home from her tiring road trip across the country. She receives a message from the emergency system: "Train crash 25 miles North from Downtown releases harmful gases. If you could be at risk, report to local authorities soon." Karen realizes that she was driving through the area about an hour ago.
Hurricane	Hurricanes are characterized by strong winds and thunderstorms around a low-pressure center. Apart from the terrible devastation to life brought by the hurricane itself, those fortunate to survive also have to deal with floods, damaged buildings, loss of loved ones, and other consequences. Keith has returned back home from vacation. He is aware that a major hurricane has caused significant structural damage to parts of the city. He receives a message from the emergency system: "Hurricane affects several parts of the city. Residents are advised to proceed with caution while entering an affected property." Keith reaches his home. He tries to enter and hears unusual squeaking sounds.
Tornado	Tornadoes are rapidly rotating columns of air which critically damage anything in their path. Though advanced tornado detection and warning systems have greatly improved odds of evacuating and so avoiding the direct impact of a tornado, special safety measures are required to ensure safety in its aftermath. Ellen has returned back home from vacation, only to learn that a local tornado has caused significant structural damage to parts of the city. She receives a message from the emergency system: "PDS Tornado Warning now cleared. Residents are advised to proceed with extreme caution while entering an affected property." Ellen drives to her home in Jefferson County, one of the affected areas. She tries to enter her home and hears unusual squeaking sounds.
Influenza	Influenza, caused by the influenza virus, is one of the most commonly occurring infectious diseases. Common symptoms include fever, sore throat, coughing, and sneezing, with the latter two responsible for its spread through air. As the virus is known to evolve rapidly, yearly vaccinations against influenza are available. Despite these efforts, influenza is known to spread around the world in annual outbreaks. Kathy was working in the office one day when she received a message from the emergency system: "Influenza outbreak reported in the city. If you are coughing, sneezing, or having a fever, see a healthcare provider soon." Kathy had been coughing for a couple of days.

Salmonella	Salmonella outbreaks are a common form of food-based bacterial infections in the US. People infected with salmonella develop diarrhea and stomach cramps up to 3 days after being exposed to the bacteria. While some people can survive without treatment, the illness could last up to 7 days. In several cases, patients are hospitalized due to severe diarrhea. David was waiting for a bus to get back home from college. He received a message from the emergency system: "Salmonella outbreak reported in the city. See a healthcare provider if you have eaten in the cafeteria & are facing discomfort." David has had mild headaches for a couple of days.

Appendix B: Summary statistics of responses obtained for various questions.

Table 2. Example of a Full-Page Table (Source)									
Construct	**Question**	**EQ**	**TN**	**IF**	**WF**	**ME**	**RA**	**HC**	**SM**
SN									
MCR	Given the scenario above, I would care what ---- think I should do.								
MCR1	my family	5.51 1.57	5.69 1.40	5.36 1.31	5.48 1.20	5.27 1.64	5.21 1.45	4.96 1.53	5.32 1.49
MCR 2	my friends	4.74 1.75	5.4 1.38	5.21 1.42	4.98 1.53	5.00 1.56	4.96 1.55	4.59 1.67	4.68 1.44
MCR 3	my colleagues	4.07 1.84	4.42 1.65	5.21 1.45	4.43 1.71	4.69 1.67	4.61 1.71	3.44 1.5	4.39 1.59
MCR 4	my community	4.47 1.86	4.76 1.68	5.08 1.66	4.75 1.72	4.97 1.63	5.14 1.58	4.33 1.57	4.21 1.57
MCR 5	other people important to me	5.09 1.66	5.47 1.50	5.36 1.37	5.28 1.52	5.36 1.56	5.32 1.36	4.93 1.57	5.14 1.30
NB	Given the scenario above ----- would want me to comply immediately with the alert.								
NB1	my family	6.29 0.92	6.31 0.88	5.74 1.19	6.05 1.36	6.27 1.15	6.21 0.99	5.78 1.01	5.75 1.32
NB2	my friends	5.93 1.24	6.24 0.74	5.52 1.54	5.96 1.03	6.48 0.71	5.96 1.29	5.63 1.01	5.36 1.28
NB3	my colleagues	5.53 1.32	5.71 1.20	5.74 1.41	5.53 1.32	6.36 0.90	5.79 1.32	5.15 1.06	4.86 1.38
NB4	my community	5.9 1.12	5.92 1.03	5.64 1.39	5.88 1.16	6.55 0.62	6.07 0.90	5.56 0.97	4.82 1.47
NB5	other people important to me	6.05 0.99	6.31 0.72	5.64 1.35	6.10 0.84	6.52 0.67	5.86 0.97	5.67 1.04	5.11 1.42
PST	The scenario above could …								
PST1	have a severe impact on my safety.	6.42 0.98	6.31 0.96	5.38 1.21	5.85 1.31	5.88 1.32	5.89 1.42	5.93 1.11	4.93 1.25
PST2	get me hurt or injured.	6.03 1.28	6.31 0.90	4.59 1.57	6.03 0.95	5.30 1.63	5.82 1.36	6.00 1.07	5.07 1.18
PFT	The scenario is likely to …								
PFT1	have a severe impact on my finances	5.00 1.25	5.83 1.34	3.62 1.60	4.65 1.79	4.09 1.63	4.36 1.70	5.74 1.38	4.18 1.42
PFT2	cause me monetary loss	5.86 1.67	6.02 0.89	3.79 1.60	4.78 1.62	4.27 1.46	4.11 1.77	5.74 1.29	4.64 1.59
IQT	Given the above scenario, how much would you agree with the following statements regarding								

IRT 1	The alert would be relevant to me.	6.18 1.13	6.35 0.93	5.77 1.16	5.95 1.13	6.06 1.09	5.68 1.47	6.00 0.73	4.93 1.51
IRT 2	The alert would be sent only when necessary.	5.74 1.21	5.58 1.36	5.03 1.79	5.30 1.65	6.86 1.36	5.32 1.72	5.22 1.05	4.93 1.51
IRT 3	The alert would be important for me to make decisions about the situation.	6.42 0.72	6.24 0.96	5.74 1.29	6.35 1.00	5.03 0.70	6.00 1.39	6.04 0.85	5.75 1.00
IAT 1	I can act on the information that I received in the alert.	6.42 0.84	6.09 1.01	5.92 1.31	6.15 1.29	6.45 0.67	6.18 0.94	5.82 0.85	5.96 0.84
IAT 2	If I follow the instructions in the alert, I will be protected.	5.73 1.28	5.02 1.27	4.62 1.21	5.28 1.38	5.18 0.98	5.07 1.12	4.85 1.10	4.57 1.03
IAT 3	The directions in the alert will help me plan my next step	6.01 1.16	5.53 1.32	5.67 1.08	5.95 1.06	6.18 0.81	5.89 0.92	5.52 1.09	5.39 0.79
ICT 1	The timing of the notification will be appropriate	5.86 1.18	5.89 1.01	5.41 1.33	5.48 1.52	6.01 0.79	5.32 1.56	5.48 1.01	5.50 0.92
ICT 2	The messages that I receive will convey the urgency for taking action	5.99 1.11	5.82 1.28	5.64 1.11	5.85 1.44	5.97 0.98	5.93 1.27	5.33 1.00	5.32 0.86
ICT 3	The messages that I receive will convey the severity of the incident	5.93 1.18	5.67 1.20	5.38 1.18	5.53 1.36	5.82 0.98	5.82 1.09	5.56 1.12	5.36 0.95
PE	Have you ever experienced any such emergency before? Please describe your scenario.								
ENS-Message Compliance Intention	When you receive an emergency notification asking you to take an action, what are you likely to do first? Please select only one option. 1. Comply immediately 2. Verify first and then comply 3. Ignore								
(EQ – Earthquake, TN – Tornado, IF – Influenza, WF – Wildfire, ME – Measles, RA – Rail accident, HC – Hurricane, SM – Salmonella, MCR – Motivation to comply with referent, NB – Normative belief, IRT – Information relevance trust, IAT – Information actionability trust, ICT – Information criticality trust; Mean and standard deviation reported)									

Appendix C: Common method bias.

Similar to the original study, common method bias was a concern since all data were self-reported and obtained from the same instrument (survey). Statistical tests (Harman's single factor test and marker-variable analysis) and procedural remedies were performed (as in the original study) to address this issue.

Harman's Single Factor test (Podsakoff et al., 2003) was conducted for all seven cases (All scenarios, Slow, Fast, Limited, Wide, Rare, Common). Single-factor analysis was performed to ensure that no factor contributes to more than 50% of the total variance, as recommended (Nov & Ye 2008; Indushobha et al., 2010). Our results indicate that the greatest variances attributed to a single factor were 34.38%, 36.39%, 33.88%, 38.65%, 32.17%, 34.50%, and 34.34% for the seven cases. The second factor accounted for 11.46%, 12.89%, 10.51%, 9.46%, 13.26%, 11.63%, and 11.95% of the variance. Thus, common method bias was not a threat to the study as per Harman's single factor test.

Marker-variable analysis (Lindell & Whitney, 2001) was also performed to identify common method bias. Since no explicit marker variable was collected during the survey, "Past experience" was used as a marker variable. The results of correlations between past experience and the other variables are summarized in Table C1. Correlations observed between the marker variable and other measured variables were not very high and so common method bias was not considered as a threat.

Table C3. Correlation between marker variable and other variables				
Scenario	**SN**	**PFT**	**PST**	**IQT**
All scenarios	0.10	0.03	0.01	-0.03
Slow	0.12	0.05	-0.04	-0.06
Fast	0.09	0.03	-0.01	-0.03
Limited	0.09	0.07	0.01	-0.08
Wide	0.09	0.07	0.01	-0.08
Rare	0.15	-0.04	0.08	0.04
Common	0.09	0.06	0.01	0.02

Procedural remedies used in the original study were also adopted in the current study. Participants never revealed their names or other identity-related information during the survey and were given the option to leave the survey at any point if they didn't wish to complete it. They were also assured that all data collected would be used for research purposes only and would be handled confidentially. Thus, variance introduced by a participant's desire to appear socially acceptable or to conform to perceived societal norms was reduced.

Appendix D: Analysis of verify-first dataset.

An analysis similar to that conducted on the unabridged dataset was conducted on the verify-first dataset. Correlation among variables were computed for the entire dataset and for categories of scenarios. The results are summarized in Table D1. Few instances (2 instead of 4) had a correlation value greater than 0.60. VIF values for these variables were below three, and so no variables were discarded from the mixed-effects analysis.

Table D1. Correlation among the variables for different scenarios (verify-first dataset)					
Scenario		**IQT**	**PFT**	**PST**	**SN**
All scenarios	IQT	1.000			
	PFT	0.251**	1.000		
	PST	0.559**	0.565**	1.000	
	SN	0.411**	0.173**	0.347**	1.000
Slow	IQT	1.000			
	PFT	0.173	1.000		
	PST	0.597**	0.494**	1.000	
	SN	0.502**	0.105	0.344**	1.000
Fast	IQT	1.000			
	PFT	0.247**	1.000		
	PST	0.521**	0.454**	1.000	
	SN	0.340**	0.241**	0.402**	1.000
Limited	IQT	1.000			
	PFT	0.267**	1.000		
	PST	**0.615****	0.546**	1.000	
	SN	0.473**	0.272**	0.556**	1.000
Wide	IQT	1.000			
	PFT	0.244**	1.000		
	PST	0.559**	0.490**	1.000	

	SN	0.364**	0.097	0.216**	1.000
Rare	IQT	1.000			
	PFT	0.146	1.000		
	PST	0.592**	0.263**	1.000	
	SN	0.429**	0.074	0.297**	1.000
Common	IQT	1.000			
	PFT	0.291**	1.000		
	PST	0.503**	**0.671****	1.000	
	SN	0.422**	0.155	0.405**	1.000
(* indicates p < 0.05, ** indicates p < 0.01)					

Table D2 shows the summarized results of the verify-first dataset at the category level. As mentioned earlier, the results obtained are similar to those in the unabridged dataset for the development speed category. Results vary in the categories of area affected and frequency of occurrence. Table D3 shows the detailed results. The models obtained for the verify-first dataset perform better (have higher correct classification rate) than those in the unabridged dataset, except for the models for the scenarios affecting wide areas and those which occur commonly.

Table D2. Summarized results of analyses for categories of verify-first dataset					
	Development speed		Area affected		Frequency of occurrence	
	Slow	Fast	Limited	Wide	Rare	Common
Past Experience	N/S	N/S	N/S	N/S	N/S	N/S
SN (H1)	**Positive**	N/S	**Positive**	N/S	N/S	N/S
PST (H2)	N/S	N/S	N/S	N/S	N/S	N/S
PFT (H3)	N/S	N/S	N/S	N/S	N/S	N/S
IQT (H4)	N/S	**Positive**	**Positive**	**Positive**	**Positive**	**Positive**

Table D3. Detailed results of mixed-effects analysis for verify-first dataset and its categories								
			Development rate		Area affected		Frequency of occurrence	
Variables	Statistics	Overall	Slow	Fast	Limited	Wide	Rare	Common
Past Experience	β	0.076	0.103	0.011	0.050	0.061	0.152	0.049
	SE	0.072	0.128	0.083	0.107	0.106	0.143	0.105
Subjective Norm	β	0.035	0.110*	0.038	0.167**	-0.016	0.009	0.099
	SE	0.032	0.051	0.040	0.049	0.044	0.043	0.050
Perceived Safety Threat	β	0.027	0.033	0.046	-0.059	0.012	0.025	-0.043
	SE	0.037	0.060	0.042	0.055	0.053	0.048	0.067
Perceived Financial Threat	β	-0.033	-0.029	-0.048	0.024	-0.051	-0.062	0.038
	SE	0.032	0.053	0.038	0.043	0.045	0.039	0.057
Information Quality Trust	β	0.144**	0.071	0.124**	0.113*	0.182**	0.163**	0.112*
	SE	0.033	0.570	0.039	0.048	0.049	0.049	0.052
Participant (Intercept)	Variance	0.144	0.165	0.163	0.156	0.112	0.139	0.113
	Std. Dev.	0.380	0.406	0.403	0.395	0.334	0.373	0.336
-2LL		335.418	146.181	211.801	161.283	208.095	203.440	173.370

Response ratio (Comply immediately: Verify then comply)	149:119	51:56	98:63	72:51	77:68	88:60	61:59
Immediate Compliance Rate	55.59	47.66	60.87	58.54	53.10	59.45	50.83
(* indicates p<0.05, ** indicates p< 0.01)							

About the Authors

Manasvi Kumar is a third-year doctoral student in the Department of Management Information Systems at the University of Arizona. He received his master's degree in Business Analytics from UIC, Chicago, USA in 2018 and a dual-degree from IIT Kharagpur, Kharagpur, India in 2016. His primary research interests include deception detection and human-computer interaction.

Gondy Leroy, PhD, is Professor in MIS at the University of Arizona. Her research focuses on natural language processing and applied machine learning. She has won grants from NIH, AHRQ, NSF, Microsoft Research and several foundations, totaling more than $2.4M as principal investigator. She earned a combined BS and MS (1996) in cognitive, experimental psychology from the Catholic University of Leuven, (1996) and a MIS (1999) and PhD (2003) in management information systems from the University of Arizona. She serves on the editorial board of the Journal of Database Management, International Journal of Social and Organizational Dynamics in IT, Health Systems, Journal of Business Analytics, and co-chairs several sessions, tracks, workshops, and conferences focusing on design science and healthcare IT. She is the author of the book "Designing User Studies in Informatics (Springer, 2011). Finally, she is an active contributor to increasing the diversity and inclusion in computing and founded and leads the "Tomorrow's Leaders Equipped for Diversity" program at the University of Arizona's Eller School of Management.