Opportunities, Options and Enhancements for the Wireless Emergency Alerting Service

December 2015

Homeland Security

Science and Technology

Carnegie Mellon University Team

Principal Investigators

Hakan Erdogmus
Martin Griss
Bob Iannucci

Graduate Students
Sumeet Kumar
João Diogo Falcão
Abhinav Jauhri

Maxim Kovalev

OPPORTUNITIES, OPTIONS AND ENHANCEMENTS FOR THE WIRELESS EMERGENCY ALERTING SERVICE

Prepared for

Department of Homeland Security

Science and Technology Directorate, First Responders Group to support the Wireless Emergency Alerts (WEA) Research, Development, Test and Evaluation Program

ACKNOWLEDGEMENTS

We would like to thank our sponsor, the Department of Homeland Security, Science and Technology Directorate, for supporting this research. We are in particular indebted to program manager Denis Gusty and his team for their direction and advice. Emergency services experts Mike Gerber, Ken Dueker and Mike Prince deserve special thanks for devoting time, providing advice and contributing ideas. Many thanks to Professor Jelena Kovacevic, the head of Carnegie Mellon University's Electrical and Computer Engineering (ECE) Department, who provided financial and moral support, sponsoring one of the kickoff events. Several CMU Silicon Valley (SV) and CMU- ECE support staff helped with logistics, accounting and financial management, which we greatly appreciate. We would like to acknowledge the invaluable involvement of over 250 volunteers who participated in the trials and interviews, without whom this research would not have been possible. They include CMU students and staff, faculty, Software Engineering Institute (SEI) staff, National Aeronautics and Space Administration (NASA) Ames community, Silicon Valley Certified Emergency Response (CERT), Amateur Radio Emergency Service (ARES), and Red Cross volunteers, other members of the Silicon Valley first responders community, and close to one hundred alert originators nationwide who responded to our online survey or who accepted to be interviewed. The Silicon Valley first responder community has in particular been extremely supportive of this research. AT&T donated some of the smart phones used in one of the trials and we are grateful for their donation.

Besides the report authors, the WEA+ team at Carnegie Mellon University included at various points in time graduate students Joel Krebs, Behrouz Rabiee, Harsh Alkutkar, Harsh Vinayak, Pallavi Mathane and Ervin Teng, faculty member Cécile Péraire, undergraduate intern Micheale Wong, former SEI technical staff member Joseph Elm, and subcontractors Mirjana Spasojevic and Rachel Hinman. Ngoc Ho and Sylvia Arifin of CMU-SV and Sheryl Benicky of CMU-ECE helped with research administration, project finances and human resources.

For information about this publication, including the corresponding
CMU-SV technical report with the full set of supporting material and appendices,
contact:

Carnegie Mellon University, Silicon Valley
Department of Electrical and Computer Engineering
NASA Ames Research Park, Bldg. 23 (MS 23-11)
P.O. Box 1 Moffett Field, CA 94035-0001

CMU-SV Publication Number: Technical Report CMU-SV-15-001

1. TABLE OF CONTENTS

List of Appendices

List of Figures

LIST OF TABLES

Executive Summary

Deployed for the first time in 2012, the Wireless Emergency Alert (WEA) service is part of the Federal Emergency Management Agency (FEMA) Integrated Public Alert and Warning System (IPAWS), providing an additional dissemination path for alert and warning messages. Authorized officials can send 90-character alerts to the public on WEA-capable and enabled mobile devices via the Short Message Service Cell Broadcast (SMSCB) protocol, a one-to-many channel for sending short text messages. The alerts are geographically targeted and a single alert is sent to cell towers in the targeted area. The alert is then delivered to all mobile subscribers covered by those cell towers. This form of targeting is coarse and does not make provision for subscriber preference, subscriber history or anticipated future movements of subscribers. The inability to provide fine-grained targeting, combined with text-based short messages' limitations in delivering adequate information and/or actionable advice, is frequently cited among factors causing citizens to opt out of the WEA service or ignore alert messages, thereby reducing the service's effectiveness.

The primary goals of this research are to gain insight into WEA adoption and acceptance issues, in particular with respect to perceived poor public response to alert messages, and to develop and test strategies for overcoming these issues within the framework of the current WEA service architecture. The methods used included interviews, surveys, social media analysis and controlled experiments.

Research Components and Methods

1. **Literature survey, interviews and surveys.** We conducted a comprehensive Alert Originator Requirements Study (AORS) to build our understanding of how alert originators (AOs) view and use the WEA service and to corroborate or refute standing assumptions and anecdotal evidence. The study consisted of three main components: a literature review, in-depth interviews with thirteen representative AOs and an online survey of 88 AOs. The interview and survey questions addressed several central topics, including: whether and how the AOs were using WEA, and what, if any, barriers they encountered in adopting WEA; the need and perceived feasibility of more precise geographic targeting; issues related to the relevance of WEA to and its acceptance by the general public; the need for regular testing of the WEA service; WEA message authoring; and the need for better education of both AOs and the general public for WEA to be a primary alerting mechanism. We also sought to uncover new requirements based on current experience with the service and opportunities afforded by smart phone capabilities, connections to social media and inclusion of links and rich media in alerts.
2. **Social media analysis.** We explored a potential connection between WEA and social media in the context of severe weather events to follow up on a central AORS theme. Selected Twitter feeds originating from areas affected by weather events were analyzed to see if there were significant correlations between the issuance of WEA severe weather alerts and increased discussions about weather in the same geographic region. Our hypothesis: WEA weather alerts trigger spikes in local Twitter traffic.
3. **Controlled experiments**. Using findings from the AORS, we explored both the usefulness and applicability of a range of enhancements and extensions to the WEA service. An important enabler for this line of investigation is the growing capability of mass-market smart phones. Computing capabilities and location awareness in smart phones hold out the possibility of personalized alert targeting that could lead to better perceived relevance and, indirectly, greater public acceptance of WEA. As the basis for making decisions of which alerts an individual user would find valuable, smart phones can securely take advantage of location history, schedule, preferences and other aspects of user context. The current WEA service's purely network-based targeting has no comparable mechanisms. Motivated by this gap, we developed a prototype of an enhanced WEA service including an alert origination system, a message delivery system and a smart phone application for processing delivered

alerts and presenting them to the recipients. We then conducted two week-long Public Usability Trials (PUT) involving three controlled experiments. The experiments used real subjects and simulated emergency scenarios. Throughout the simulated emergency scenarios, we assessed the effectiveness of new system features by measuring and comparing the responses of the participants to alerts issued in real time.

4. **Architectural studies.** While developing the WEA enhancements and extensions tested during the controlled experiments, we examined the feasibility and conceptual costs of these new features within the current WEA framework. We developed architectural alternatives in this context, integrated the ones best positioned to support the new features into the experimental infrastructure and tested them as part of (3) to assess their feasibility.

The central findings of the above investigations and associated recommendations are summarized below.

Accomplishments, Results and Recommendations

1. **(a) Outreach and education are necessary for WEA's success**: Our interviews and surveys revealed the need for AOs and the public to be better educated about the WEA service. Some AOs in our study did not use WEA – when it might have been appropriate to do so – simply because they were under-educated about it and, as such, saw its use as a potential risk. The phenomenon of the public opting out may be attributable to the fact that they do not understand the kinds of messages which they then might miss. We recommend an outreach program aimed at significantly improving the understanding of how WEA uniquely serves the public interest.

   (b) Effective education must be coupled with testing and testability: AOs further highlighted the need for regular systematic testing of WEA. Not unlike Civil Defense, the Emergency Broadcast Systems and the Emergency Alert System AOs reported that to be effective in times of need, the WEA service needs periodic testing with sufficient participation from the public. This would serve the purpose of giving AOs practice in using the system as well as showing the public that the service provides a unique and valuable means of receiving alerts. We recommend the creation of a systematic WEA testing program.

   The WEA service has **unique** aspects that make it technically more complex than some other public warning and preparedness systems. Its use requires careful crafting of messages. The capabilities of end-user devices (smart phones) and the evolution of the delivery medium (cellular technology) present new and changing opportunities for improvements. Our research demonstrated the WEA service would do well to evolve into a richer communication mechanism and improvements, including potential new features, can be tested with volunteers using an evidence-based approach. As such, there is a need for a developmental platform, not unlike a firefighting or police training facility, at which AOs, government agencies and emergency organizations can safely pre-test and experiment with new WEA concepts, features or enhancements before these improvements are deployed generally. We propose the creation of a national-level WEA testbed to serve this need.

2. **Deep integration of location-based context materially improves WEA's value**: The AORS indicated that AOs believe increased geo-targeting resolution will significantly improve the effectiveness of WEA. Motivated by this finding, we developed techniques for compressing boundary polygons of alerts' geographical targets and WEA message texts to enable the smart phone, rather than the network, to make the *alert-or-not* decision. Our on-phone decision-making software further supports triggering based not just on position, but also on position history. The PUT component of our research demonstrated strong evidence of acceptance of and value for

these new features. We recommend WEA alert creation tools be modified to allow embedding of boundary polygons into WEA messages. We further recommend that smart phone manufacturers build flexible mechanisms into future phones to take advantage of boundary polygon information and other user context information available within the phone in making the alert delivery decision.

3. **WEA's relationship to relevant social networking channels should be re-thought and clarified**: When SMSCB was created, there was no comparable pervasive alerting technology. Nowadays, smart phone users rely on a wide range of social networking services for access to news, and WEA is simply one of a number of information channels available to the public. The use of services such as Nextdoor.com by police for two-way communication with residents demonstrates that public expectations have changed, possibly in fundamental ways at odds with the current architecture of WEA. When asked about the role WEA should play in a world where social networks have become pervasive, AOs well-versed in social networks see WEA playing the role of an *alarm bell*, rather than the role of the sole source of authoritative alert messages. They identify WEA's short text message basis, targeting limitations and lack of provision for rich media as reasons why WEA as it exists today is unlikely be the main source of alert information sought by the public. Our preliminary analysis of social networking activity relating to severe weather events showed no evidence of any correlation between the timing of weather-related WEA messages and Twitter posts about the underlying weather events, suggesting that in its currently isolated form, WEA probably does not play a significant role in leveraging social networks to widen its reach in the context of weather alerts, the context in which it is most often used.

   Based on these findings, it is our opinion that WEA must be re-formulated as a complement to existing and future social networks, and WEA's fundamental purpose (whether it should serve as an alarm bell, enhanced service with social-network-like capabilities, rich media channel or a combination of thereof) must be re-established.

   This opinion raises a number of important challenges about most aspects of WEA. It is our recommendation that a major study of how WEA should evolve to fit into the communication expectations of current and future citizens is justified.

4. **Rich-media integration into WEA is a question of *how*, not *if***: The AORS provided support for the integration of rich media (such as use of photos, maps and carefully typeset and laid-out text as one would find in a well-crafted web page) into the WEA service. PUT results indicated that integrating maps showing the alert region and recipient's location into WEA alerts were perceived as highly desirable. We believe that the call to include rich media content in alert messages arises from the fundamentals of (a) widespread use of smart phones and (b) the pervasiveness of the World Wide Web and the ways in which the Internet sets the standard for how information is conveyed. When SMSCB was selected as the WEA information transport mechanism, cellular networks did not support broadcast of rich media. Since then, advances in cellular network architecture (such as the 4th Generation Long Term Evolution, or 4G LTE, broadcast and in-network content caching) have made rich media broadcast possible. Similarly, inclusion of maps will be facilitated by the availability of pre-cached, built-in maps in future generations of smart phones. In light of these anticipated developments, and both AORS and PUT findings that indicate perceived value for content-based enhancements, we recommend a re-consideration of WEA at the level of network standards bodies and builders of WEA smart phone software to support different content forms.

5. **WEA will benefit from a transition from a focus on *alerting* to a focus on *awareness***: WEA's fundamental nature as a 90-character text message broadcasting service may work well in situations where the descriptions of the emergency events and the recommended preparedness actions are simple and relatively unchanging. For situations which unfold over time and in which instructions to the public may be revised in the span of minutes to hours with many different, possibly conflicting alerts and updates issued (e.g., in the case of an earthquake causing bridges to collapse, triggering fires or release of hazardous materials, or requiring management of changing evacuation protocols), WEA's means of presenting information to the subscriber may not be well suited, however. Errors from interpreting individual text-based alert messages and updates out of their original order may lead to serious consequences. We demonstrated through an experiment conducted in the context of the PUT component of our research that providing software on smart phones to *digest* sequences and sets of related WEA messages and to present the digested information as a *situational-awareness* view resulted in significantly better understandability compared to the standard WEA presentation of alerts on the phone. We refer to this new way of structuring and viewing streams of alert messages as a change in WEA service from a focus on *alerting* (sending many messages) to a focus on *awareness* (assisting the user by digesting what has been sent into a comprehensive, up-to-date view). WEA's implementation today does not support such digesting. To address this, we also developed the means by which this digesting capability can be retro-fitted to the current WEA architecture.

1. INTRODUCTION

The Wireless Emergency Alerting (WEA)[1] service is part of the Federal Emergency Management Agency (FEMA) Integrated Public Alert and Warning System (IPAWS), providing an additional dissemination path for alert and warning messages (SEI 2014). Authorized officials can send short text alerts to the public on WEA-capable mobile devices via the Short Message Service Cell Broadcast (SMSCB) protocol, a one-to-many channel for 90-character text messages, wherein messages are geographically targeted and sent to cell towers covering an affected area, and are subsequently delivered to mobile subscribers.

This report presents the results of an extensive investigation into the WEA service from two main perspectives: the Alert Originator (AO) and the public. The primary goals of the study are to gain insight into WEA adoption and acceptance issues, in particular with respect to perceived poor public response to alert messages, and to develop and test strategies for overcoming these issues within the framework of the current WEA service architecture. The work was organized to reflect these points of view.

In the first phase of our work, to deepen our understanding of the issues confronting AOs as they consider adoption and use of WEA in addition to or instead of other alerting mechanisms, we designed and conducted two investigations with the AOs. We call this phase the AO Requirements Study (AORS). The AO investigations helped us to validate or refute standing assumptions about the WEA service, identify its strengths and shortcomings and solicit recommendations on how to improve its adoption, effectiveness and relevance. Structurally, we organized the AORS work into two components:

1. In-depth interviews with 13 carefully-chosen emergency services personnel, each with high level of experience and familiarity with emergency communications; and
2. A larger-scale, 30-question online survey of the wider emergency alert community, to which we received 88 responses.

In the second phase of the work, two **Public Usability Trials** (PUT) explored potential enhancements to the WEA service. In this phase, informed by the AORS, we proposed, developed and evaluated a variety of extensions to WEA aimed at addressing the identified shortcomings. Importantly, the evaluations included the development of a powerful and flexible testbed and experimental framework, allowing us to engage more than 225 subjects in simulated-emergency situations and equipping them with futuristic WEA implementations. By measuring and assessing the subjects' abilities to assimilate information about fictitious emergencies through WEA — both with and without the extended features — we created a mechanism by which we could quantify the perceived *value* of each feature. We conducted three experiments using this approach over the two trials. Lessons learned during the initial trial served as a basis for extending both WEA and the testbed framework itself for the second trial. Such live testing of WEA in a realistic setting with human users is a unique aspect of our work, and, to our knowledge, constitutes the first WEA study of its kind.

As a complement, we also performed architectural studies to examine the *feasibility and conceptual costs* of the new features in the current WEA framework, to determine what could be done with simple 90-character text messages as a data transport mechanism — but extending the manner in which the 90 characters can be encoded and interpreted. Many of the resulting architectural improvements were integrated into the testbed framework to support the new WEA features that leveraged them.

---

[1] WEA was formerly known as the Commercial Mobile Alerting System (CMAS).

1.1. Starting Assumptions

WEA's essential purpose can be thought of as delivering the right alert messages to the right people at the right time via mobile phones subject to (perceived or real) constraints of how such alerting can be done using cellular networks. Herein is a central challenge: *targeting*. How do we know a priority, and on a mass scale, which alerts will be relevant for which people, and how do we build an efficient system for delivering alerts accordingly? A second central challenge is *content* — what, exactly, is the right emergency alert? How can the message — or messages — be conveyed to the recipients in ways to maximize understanding and minimize opportunities for error in the midst of a complex, evolving and stressful situation?

The current WEA service simplifies the problem by using an individual's then-current location as a valid proxy for his/her interest in an alert with some geographic specificity.[2] If there is a wildfire in the area where a person's daughter's school is located, and he or she is currently 100 miles away, the system might well infer that the person is not interested when, in fact, he or she is. We may call this misconception the *location-proxy fallacy*. Another consequence is the potential of sending irrelevant messages to people in the area, but otherwise not affected or interested — resulting in the user turning WEA off or ignoring incoming alerts. We call this response the *opt-out problem*.

The current WEA service also constrains the solution by adhering to the limitations of 2G cellular networks as inherent in how alerts are constructed (as simple, unstructured text strings), disseminated (using SMSCB) and interpreted (text strings shown as disembodied pop-up alerts). Yet, a cursory examination of how large-scale disasters play out, and how messaging takes place in those situations, reveals the need to periodically *update* messages as situations unfold (effectively, being able to erase and replace obsolete prior alerts — lest the population act on out-of-date information). It reveals the need to represent the information not just as a linear sequence of messages, but rather as an *interconnected tree* of alerts (with each branch being subject to a determination of user interest). It also reveals the need to convey messages using *rich media*. We refer to the misconceptions limiting the current WEA solution to unstructured, unrelated streams of short messages as the *short-message fallacy*.

In both of these cases, we use the adverb *fundamentally* because the location proxy fallacy and the short message fallacy are so deeply built into WEA that few would seek to overturn these assumptions. Yet overturn them we must.

1.2. Technological Advances Impacting WEA Expectations

From its inception, the WEA service has been based on cell broadcast technology, more specifically the SMSCB protocol (3GPP 1999). The SMSCB was first demonstrated in 1997. It grew out of a need to be able to send messages to large populations without suffering the time penalties of issuing individual text messages to each subscriber. Such a system is particularly useful during emergencies when traffic on the wireless networks is especially high or when network capacity has been degraded.

Since the initial demonstration of SMSCB, network capacities have improved dramatically (in access networks, cell backhaul and the core networks), new protocols such as Long Term Evolution (LTE) broadcast have emerged, and *user expectations have shifted* from text messaging to rich messaging.

---

[2] This is similar to strengths and weaknesses of sirens as alerting devices — almost all people in the area will hear the siren without having to sign-up (unless underground or hearing impaired), but AOs have limited control over who receives the alert.

In addition to the evolution of network capabilities, society has benefitted from the emergence and now-pervasive use of smart phone technologies. When SMS cell broadcast was created, phones were little more than devices for making voice calls and handling text messages. Now, smart phones have capabilities exceeding those of personal computers from a decade ago.

Considering how technologies in both the network and the phone have changed dramatically since the inception of SMSCB, we must ask if the rising tide of expectations compels, and the technology enables, a significant re-thinking of WEA.

1.3. Key Concepts

Based on the starting assumptions and the technological advances that potentially reset AOs' and the public's expectations of WEA, our work proposes two key concepts that may help improve the adoption and effectiveness of the WEA service: (1) exploiting the context and (2) moving beyond alerting to awareness. Each of these key concepts are expounded upon next.

Exploiting the Context

Exploiting the context is essentially inferring interest based on each user's individual context to avoid the location-proxy fallacy and to minimize opt-out. Barring the user's ability to suppress the reception of certain alert types (viewing user preference as a component of identity) and the alert's target zone (a component of location), the current WEA service does not take much advantage of either the user's context or the alert's context to make alert messages more individually relevant to the recipients, prevent over- or under-alerting and increase the effectiveness of the WEA service. The simplicity of the WEA service — while it may be advantageous from certain standpoints, such as bandwidth efficiency — largely wastes advanced capabilities and computational power of modern mobile devices.

Significant context information is available to modern smart phones and other mobile devices. If alert messages can be augmented with metadata about the context of the alert itself, the built-in WEA app on recipients' devices can combine this metadata with information available or inferable by the device about the user's own context to make a better alert delivery decision, personalize the alerts or even give customized actionable advice.

In particular, context awareness can help with geo-targeting precision. Currently, WEA messages are broadcast through cell towers in the region identified by the alert. The density and range of cell towers vary and the ranges overlap, however. Anecdotally, wireless carriers are opaque about how they map an alert region (for example specified by a polygon) to a set of cell towers. But even with more openness about the reality of how targeting mechanisms have been implemented, precise geo-targeting is simply not possible with cell broadcast technology alone. Most modern phones are equipped with a good consumer-grade Global Positioning System (GPS) receiver and WiFi capability; built-in location services can determine the geographic location of the device with reasonable accuracy by combining information from these two sources. If the geo-target specifying the alert region can somehow be embedded in the broadcast alert payload, alert messages can be filtered on the device by the built-in WEA app by checking the device location against the alert region. We refer to this client-side geo-targeting feature as geo-filtering. Geo-filtering effectively reduces the dependence on wireless carriers to control geo-targeting precision. Other customizations are also possible: for example, the threat level could be dynamically adjusted based on the recipient's proximity to the alert region.

Today's smart phones include network-based mapping capabilities and built-in maps are expected to be standard in future smart phone generations. Geo-filtering can be combined with high-information maps to give the recipient easily absorbable visual cues. This combination has the potential to meet the opt-out problem head-on and also to address the short-message fallacy. A straightforward approach would be to simply show the location of the user relative to the alert region. The proximity of the recipient to the danger zone could also be calculated, and the device may suggest the nearest exit, the best evacuation route or the best evacuation direction to the user. In such ways, location-based context information could be highly valuable.

More advanced uses of location information include location history and location prediction, which can augment geo-filtering. These enhancements address the location proxy fallacy of our starting assumptions discussed in Section 1.1. Location history involves the recipient's device keeping a record of frequently and recently visited places. This record can be used during geo-filtering in a predictive way based on the assumption that a person who visits a location frequently or has visited a location recently more than once is likely to be prefer to receive an alert affecting the vicinity of that location even if the person is outside the targeted alert region. Similarly, users' speed and direction as determined by user devices' built-in GPS could be used to predict their presence in an alert region in the near future, which could again be used during geo-filtering. This latter capability is what we call location prediction. No privacy or security concerns are raised in these instances since all processing of contextual information is performed on the client side: private or confidential information never leaves the user's device.

A user's mobile device can also store the user's preferences or profile locally. This is much more efficient, straightforward, feasible and secure than registering such information with the alert originating entity, and may help improve accessibility of the alerting system. For example, a visually impaired person may prefer to receive alert notifications via text-to-speech (TTS) or a user may opt to receive alerts in format while he or she is asleep. User's, or the user device's, identity includes real-time connectivity information, which is readily available. If alert messages can carry external links, the user device can use this connectivity information to decide whether to show a Uniform Resource Locator (URL) embedded in the alert payload depending on whether the device has Internet connection when the alert is viewed. These examples constitute ways in which the identity-related (of the user or the device) contextual information can be leveraged.

Time and activity are other components of context that may help increase the relevance and actionability of alert messages to the recipients. Many modern mobile devices, in particular new generation smart phones and wearables, have built-in gyroscopes and accelerometers as well as application programming interfaces (APIs) that allow a user's current physical activity to be inferred. For example, if a person is sleeping at home after midnight, an AMBER[3] alert may not be relevant or actionable to that user, but a similar message arriving while the person is driving or cycling would be welcome. Also a TTS notification may be preferable to the standard alert tone when the person is driving or cycling.

The above illustrates a few examples of how contextual information could be used on the users' phones to filter and customize alert messages and to determine appropriate delivery modes. The types and combinations of such contextual information are numerous, and there are many ways of leveraging the combinations. In the two trials of the PUT phase, we assessed a subset of such combinations by implementing selected WEA enhancement features based on them. These features include precise geo-targeting (via geo-filtering), location history and location prediction.

Note that in the standard WEA service, geo-targeting is achieved by SMSCB at the cell tower level. Therefore it is dependent on the mapping of the AO-specified alert region by participating wireless carriers to a subset of cell towers in their coverage areas. To gain maximum advantage of client-side context-based filtering, the broadcast boundaries of messages in the network must, necessarily, be expanded. It remains to be decided how widely alerts should be sent so as to avoid false-negatives (relevant messages that do not arrive due to network-side filtering). Our focus is on increasing the percentage of users that find a given alert relevant through adding client-side capabilities.

---

[3] America's Missing: Broadcast Emergency Response

As part of our architectural studies and in support of making phone-side geo-targeting a practical reality today, we also explored the ways in which targeting polygon information can be embedded directly in WEA messages — compactly and accurately. Results of the geo-target compression were presented in a paper by Jauhri, Griss and Erdogmus (2015), and is summarized in this report.

Moving Beyond Alerting to Awareness

Although WEA may be suitable for alerting with a single message, asking users to mentally stitch together sequences of messages related to the same event may lead to confusion and incorrect action. Avoiding such undesirable consequences implies abandoning the short-message fallacy and considering how the evolution of technologies has changed both what is possible and what is expected — for the purpose of creating awareness, not just for the sending of alerts.

Confusion and incorrect action are more likely in a complex and evolving emergency such as a large-scale earthquake. The ability to effectively deal with such scenarios motivated us to explore alternative alerting schemes subject to the constraint of preserving existing WEA machinery where possible. Our hope was to measurably reduce user confusion in the presence of complex message sequences involving multiple, possibly inter-related, incidents.

As a means for delivering individual, stand-alone text messages, WEA is bandwidth-efficient, and presentation of the resulting text strings is easily fit into the notification systems of modern smart phones, using pop-up alerts, as shown in Figure 1.1.

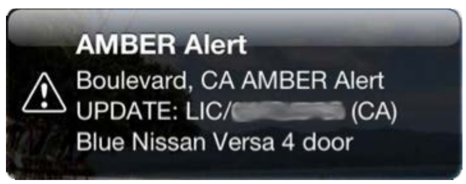


Figure 1.1. Example of an Ordinary Pop-Up WEA Message

If the need were to simply convey a single, unchanging, guaranteed-to-be-correct message per emergency situation to a large population, WEA as described would suffice. But let us consider the alternative cases: conveying more information than will fit a 90-character message, or messages that update or correct prior messages. Is it right to assume users will correctly digest these sequences and updates so as to have the most accurate information in mind when they are making potentially life-threatening decisions about where to go and what to do?

We concern ourselves here with this more complex case of using WEA as a means to guide and direct the actions of a large population during an emergency situation of sufficient complexity as to require the sending and digestion of multiple messages. These messages might span different but related hazards, each of which might change over time. With this time evolution of the emergency situation, AOs will need to send update alerts as the situation evolves; the content of any given message might augment or even contradict one or more prior messages. Also, if the region is changing as the event unfolds, some prior messages may not reach newly affected users. It is also possible for the recipients to react differently to update alerts compared to initial alerts. Then would the new information be correctly assimilated and synthesized with existing information to lead to proper action, especially as the emergency scenarios became increasingly more complex?

Consider the example of a large-scale earthquake that causes initial physical damage (e.g., collapsed bridges) followed by fires, traffic jams and secondary hazards, such as the release of toxic airborne contaminants. AOs might issue an initial message identifying the bridge collapse with guidance (e.g., *"avoid the Dumbarton Bridge and its approaches"*). As the situation unfolds and fires erupt, AOs may issue shelter-in-place messages. With the release of toxic material, shelter-in-place may be revised to call for evacuation of targeted areas. As the toxic plume moves and additional fires erupt, the correct instructions may continue to change, rendering prior instructions moot or, in the worst case, incorrect. Such a barrage of messages, particularly bearing updates and changes of strategy, require individuals to receive and digest them in time sequence, maintain a mental model of the latest instructions and be able to recall these when acting. The mental complexity of this process, combined with the high stress of the situation, can easily lead to errors of action on a wide scale.

We propose an alternative to the current WEA architecture that, we contend, could lead to fewer errors of interpretation. Specifically, we propose a revised system in which:

- Alerts are originated using the full expressive power of Common Alerting Protocol (CAP), including all the relevant metadata as opposed to just a few fields. The CAP (OASIS 2010) is a standard adopted by many AOs to comprehensively define an alert and its context, but the current WEA service only relies on CAP fields that can be used to construct the alert text and a few others related to emergency type and severity.
- CAP messages directed to WEA channels are encoded in such a way that they can be packed for and transported over the existing WEA 90-character channel. We refer to this encoding as WECAP, which we discuss in detail in a companion publication (Iannucci et al. 2015).
- Alerts so represented preserve important information such as the relationship between messages in a sequence. This change also affords much greater power in deciding how the alert can be most meaningfully presented to the user based on user-specific context.
- WECAP messages arriving at users' smart phones will be unpacked, organized into sets of sequences (based on message identifiers) and presented as a set of digests, one digest per sequence. Digests are created by playing back message sequences, recording the most recent information for each coded CAP field. If done properly, the result will be an aggregate of CAP fields, each containing the most up-to-date information.

Importantly, this scheme preserves much of the existing WEA transport mechanisms, but replaces the presentation with a new, digested format which we refer to as *Situation Digest*.

Consider the above earthquake scenario. AOs issue a first alert with instructions to avoid the impacted area. With the outbreak of fires, they issue a *shelter-in-place* directive. Then, when the toxic plume is detected, this directive is revised for the appropriate area from *shelter-in-place* to *evacuate to the south*. Evacuation-related traffic jams to the south may prompt a revision to *evacuate east — avoid southbound Highway 101*. With today's WEA, impacted individuals would receive four messages and would need to read them in the correct order to know to evacuate to the east, and some users may not receive all the messages if they move or if the affected region changes. Under the proposed revision to WEA, calling up the presentation of alerts would *only* show the most recent directive — to go east — avoiding the confusion altogether. Figure 1.2 depicts this process.

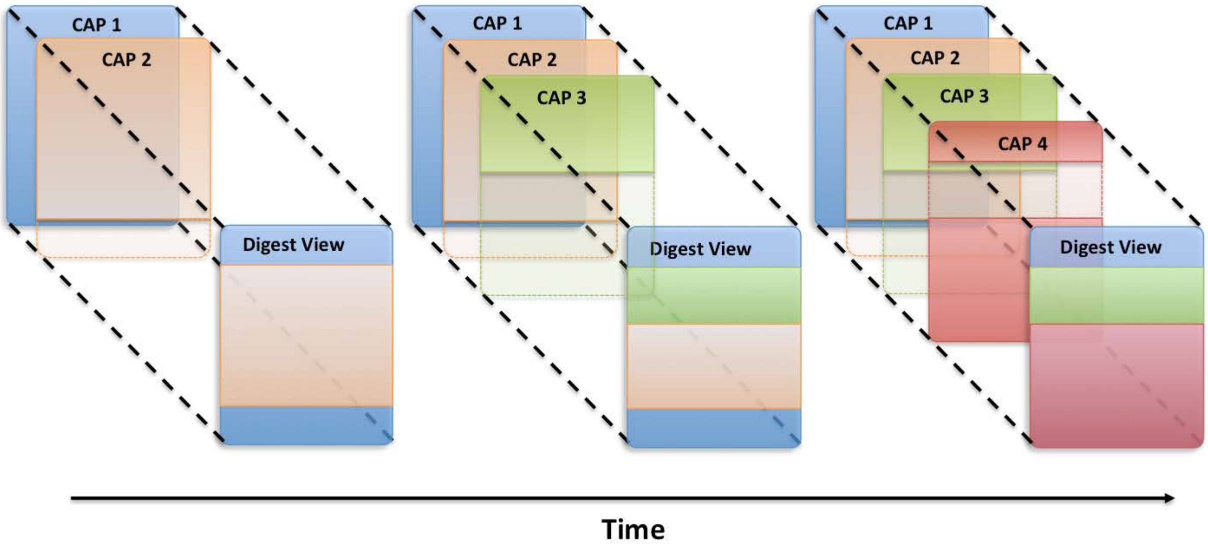


Figure 1.2. Digesting a Stream of Messages Into a Single View

1.4. Alert Originator Requirements Study: Methodology

The AORS phase aimed at understanding the AO perspective and their requirements. We needed this perspective before we could devise and test different WEA enhancements and concepts with real users. We therefore conducted a set of in-depth interviews followed by a survey. The interviews involved 13 AOs. We also received 88 responses using a 30-question online survey. The participants in these two studies were selected from a pool of more than 500 AOs across the United States.

Many of the questions we asked in the two AO studies were developed after a thorough study of prior work,[4] with the intent of validating some of the known issues as well as identifying new issues. We were interested in how elements such as message length, geo-targeting specificity and education of AOs and the public about WEA's use and purpose affected the service's adoption by AOs and the public's perception of its effectiveness and impact.

Section 3 describes the AORS process and results in detail, and Appendices A1 to A3 provide additional detail on the goals, design and findings from the interviews and survey. We analyzed the collected interview data using the Grounded Theory method (Strauss and Corbin 1998), a qualitative approach commonly used in social sciences. Most of the survey questions generated summary tables segmenting the various responses, and a few questions with free-text responses were analyzed using thematic analysis (Braun and Clarke 2006).

Key results from this phase of the work pertain to the low WEA adoption rate among AOs, need for better education for both the AOs and the public, ability to test and evaluate system effectiveness, advantages of a wide-reach channel, importance of geo-targeting specificity, disadvantages of message length limitation, inclusion of maps and rich media in alerts and integration of WEA with social media.

[4] A summary of relevant prior work is provided in Section 2.

1.5. Public Usability Trials: Methodology

As an exercise in measuring the value of proposed WEA extensions, we faced the challenge of assessing how users interpreted the alerts. We evaluated these outcomes (selectively applied, when applicable to studied feature/concept):

- Understanding: Was the alert easy to understand?
- Relevance: Was the alert relevant to the recipient given the recipient's context?
- Annoyance: Would the alert annoy the recipient in a similar real emergency situation?
- Actionability: Would the alert prompt the recipient to take protective action in a similar real emergency situation?
- Milling behavior: Would the alert the encourage recipient to seek confirmation from alternative sources?
- Adequacy: Does the alert contain sufficient information for the recipient to assess it?
- Usefulness: Would the recipient find the alert useful in a similar real emergency situation?
- Situational awareness: Does the organization, aggregation and presentation of messages in a complex emergency situation affect the recipient's awareness and understanding about the evolving nature of that emergency?

We designed the PUT to focus on: (1) simple, phone-centric enhancements that promise to improve public adoption of the WEA service and prevent opt-out and disengagement in general; and (2) a structured alerting approach that directly targets sense making with changes to alert creation, transport and presentation, and promises to reduce confusion and trigger correct action in complex emergency situations.

For the **first trial** (Trial 1), we elected to focus on alert targeting specificity and richness of content based on the user's context (location in particular) and various alert additions, respectively. Alert additions included TTS notifications, maps and external links. Context-aware features included fine-grained, precise geo-targeting with filtering on the phone, which addressed geo-targeting specificity alone, and additional filtering based on the user's location history and movement, which addressed targeting specificity with a wider range of contextual information.

For the **second trial** (Trial 2), we developed a new concept that specifically addresses situational awareness, along with a supporting alerting system and a representative user-facing feature — a digested situational awareness view — that implements the concept. Situational awareness is defined as the ability of an alert recipient to make sense of the information contained in an alert message and make correct inferences about what the underlying emergency is about, what to do next and when to act. We then studied whether this concept delivers the hypothesized benefits.

Across both trials, we prototyped, trialed and measured the perceived value of:

- Long messages,
- High-information maps,
- Phone-side precise geo-targeting (via geo-filtering),
- The use of location history and location prediction in geo-targeting,
- Enhancing message content with external links,
- Delivery of alerts using TTS, and
- Presenting message sequences and sets via a digest view rather than as a list of text messages.

2. Related Work

Several previous studies and workshops have reported on the effectiveness and adoption of the WEA service and similar alerting mechanisms, and suggested strategies for possible improvements for WEA delivery. We summarize the most relevant ones here.

2.1. Geo-targeting

In the standard WEA service, geo-targeting is achieved by SMSCB at the cell tower level. Therefore it is dependent on the mapping of the AO-specified target region by participating wireless carriers to a subset of cell towers in their coverage areas (Figure 2.1).

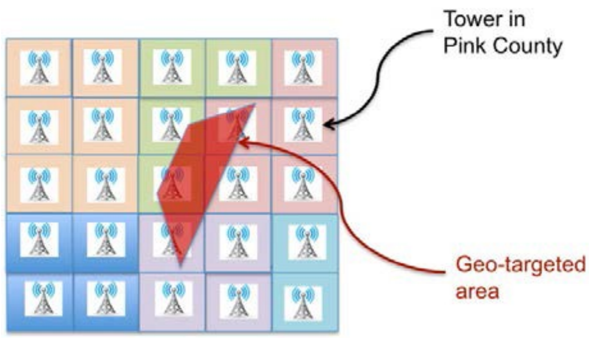


Figure 2.1. Geo-Targeting Specificity: Illustration of Mismatch between Geo-Target and Cell-Tower Coverage

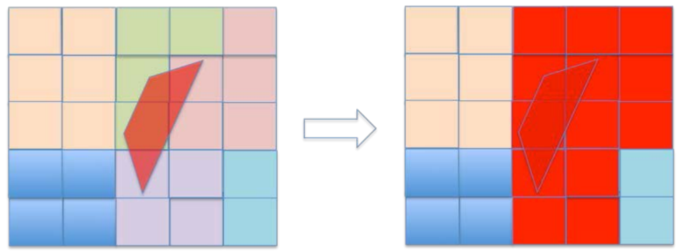
The precise, fine-grained implementation that we propose here adds client-side filtering to this method: it is not meant to replace the standard base-station-based targeting, but rather to augment it. Geo-targeting accuracy and specificity can be improved in different ways along the alert creation and delivery pathway: at the source by better prediction of the affected area, on the way by better mapping to cell towers and at the receiving end by matching recipient location with the geo-target representation embedded in the alert message, via geo-filtering. We opted for the latter approach, which is easy to implement, feasible and low cost. Geo-targeting options are compared in Figure 2.2 with respect to specificity. In each pair of images, the image on the left is the desired geo-targeted area, the same in all cases, while the image on the right shows the resulting targeted area: (a) activates too many cell towers, (b) activates only cell towers in enclosing areas, while (c) only activates cell-phones in the exact desired area via geo-filtering.

Research conducted at SEI (Stoddard et al. 2013; Woody and Ellison 2013) identifies trust, both on the part of the AOs and the public, as a key factor in the success of the WEA service. Based on an analysis of the AO trust model they developed, SEI authors determine that maximizing AOs' use of the WEA service requires maximizing three key outcomes: appropriateness, availability and effectiveness. SEI reports (SEI 2014; McGregor et al. 2014; Stoddard et al. 2013) also suggest that geographic specificity is a critical component in building trust through appropriateness.

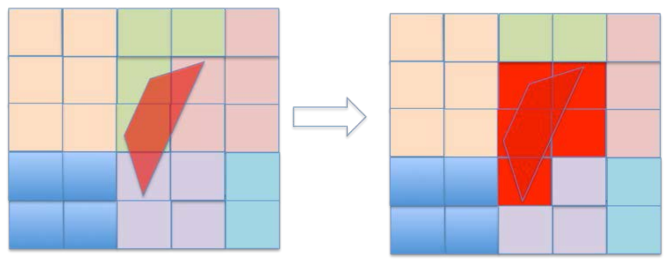
Nagele and Trainor (2012) state that being able to set an appropriate polygon size is an important factor in improving the response to alerts. This observation was confirmed by our AORS in interview and survey responses. It was sometimes expressed as the need for smaller polygons to specify a targeted geographical area, and for specified polygons not to include all of the enclosing FIPS[5]-coded regions.

---

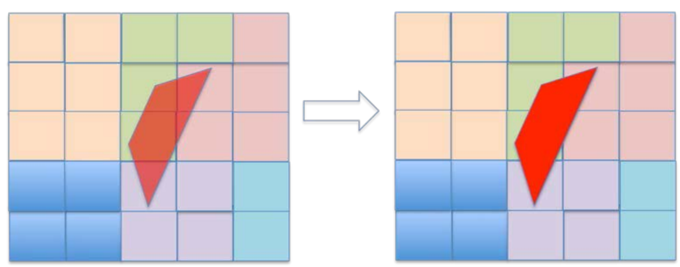
[5] Federal Information Processing Standard



(a) Poor Geo-Targeting



(b) Better Geo-Targeting



(c) Precise Geo-Targeting

Figure 2.2. Comparison of Options to Achieve Better Geo-Targeting Specificity

The importance of geo-targeting is also reiterated in several other reports. The Department of Homeland Security's (DHS) WEA service recommendations (DHS 2013) and SEI's WEA best practice recommendations (McGregor et al. 2014) conjecture that AOs will use WEA messages more extensively if alert messages can be better targeted to the size and location of the geographic region impacted by the emergency event. In particular, one of the DHS recommendations state that geo-targeting precision enhancements "would prevent missed alerts caused by geo-targeting inaccuracy and reduce over-alerting the public with irrelevant messages," adding that such outcomes would in turn "encourage more widespread adoption of WEA by emergency managers and the public."

The WEA service targeting based on the current county designations using the FIPS codes are effective in some cases, but not all. For example, in some states, counties have a huge geographic footprint and notifications of an emergency in the far corner of a county may send useless information to many who are hundreds of miles away. Conversely, in major metropolitan areas where the distances are smaller but population density is higher, current WEA geographic granularity may result in many people receiving alerts for a localized event not relevant to them.

SEI's study on WEA integration considerations (SEI 2014) contrasts the targeting approach based on the FIPS codes with a more granular polygon approach. Many AO tools support the specification of an alert area with a polygon, but the WEA messages are often issued to all FIPS codes intersected by that polygon, significantly expanding the originally targeted area. Even when commercial mobile service providers support polygon-based targeting, it is not clear to what degree they are over-approximating the specified polygons during cell-tower mapping, and how such mapping affects disparities between sparsely populated rural areas with few long-range cell towers and densely-populated areas with many short-range cell towers. While Tier-1 carriers[6] have started using more granular geo-targeting, changes in cell tower coverage and differences in geo-targeting precision related to different tower and population densities continue to maintain the status of targeting resolution as a top concern. Moreover, dynamic changes in tower power and coverage means a static understanding of coverage is not sufficient.

In our study, precise geo-targeting via geo-filtering was implemented using built-in location-based services on client devices and required the transmission of the geo-target with an alert message. Location-based services that rely on hand-held devices' GPS and Wi-Fi capabilities have been in use in mobile applications for quite some time: ubiquitous usages include navigation, local weather reports and local traffic reports. Zickurh (2013) reports the use of location-based services have increased considerably in the past years, with over 74 percent of smart phone users in 2013 stating to have taken advantage of these services for some purpose. There was simultaneously a modest decrease in the percentage of smart phone users sharing their locations with friends due to privacy concerns, however. Jungras and Watson (2008) suggest users are nevertheless inclined to forego privacy concerns if the resulting services are deemed sufficiently beneficial, for instance in public safety and security applications. Examples of use of location-based services in emergency domain is not new, with several cases having been reported as early as 2007 (Aloudat, Michael and Yan 2007). The idea of using of such services in the context of wireless emergency warning systems is relatively recent, however.

2.2. Inclusion of Maps in Alerts

Location-based services are often most appealing when combined with maps that allow the users to visualize pertinent information in the proper context and in real time. Such maps could be called *high-information*. Though maps have been frequently used in mobile applications, their use in emergency alerts delivered to mobile devices is still not very common. At the time of writing this report, a survey of emergency-warning related apps on Google PlayStore turned up very few Android-based alerting systems with a map capability (Government of Alberta 2015; EC Network 2015; Elecont Software 2015). Moreover, each of these systems appeared to have a very small user base (in the thousands). Therefore, the WEA landscape currently includes no widely used map-based service, even though previous research has pointed to potential benefits of map-based solutions.

Hagemeier-Klose and Wagner (2009) used a web-based map for flood hazard alerting and Daly et al. (2014) used device-generated maps with the recipient's location. Most recently, Hamilton et al. (2014) conducted a study that examined public response to WEA using *in-vitro* experiments, focus groups, a survey and interviews. In this study, the inclusion of a high-information map (specifying the alert region and the recipient's location) had a statistically significant and positive effect on public response outcomes including interpretation and personalization, with a potential to improve protective action-taking. Unlike the *in-vivo* experiments conducted in our trials, which relied on real-time test alerts delivered to subjects' mobile phones using a WEA emulation testbed with various enhancements, Hamilton et al.'s relied on subjects filling out a questionnaire in a laboratory setting about their response to hypothetical alerts.

---

[6] A Tier-1 carrier is a commercial wireless service provider of voice and data services with a telecommunications network solely owned and operated by the provider and having direct access to the Internet.

A recent study by Casteel and Downing (2013) contradict the above findings, however. This study, using National Weather Service (NWS) warning messages sent through NWS's iNWS messaging service, compared warnings that included a radar image of the storm to those that were purely text based. The results showed participants' understanding was not influenced by the inclusion of storm images. None of the other message content outcomes differed either with the presence or absence of the images. The authors suggest these results should be considered in evaluating similar options for the WEA service.

2.3. Miscellaneous WEA Limitations and Enhancements

The SEI study on WEA integration considerations (SEI 2014) posits that continual WEA technical improvements in an evolving infrastructure and the ubiquity of smart phones may enable novel technical solutions both for improved geo-targeting and for addressing other current limitations of WEA-based services. This potential of capable mobile devices to address such limitations was also previously alluded to in both the 2011 and 2013 editions of the workshops commissioned by the National Research Council (NRC) on current knowledge and research gaps (NRC 2011; NRC 2013), the 2013 edition of which reported that alerting systems in the future may not be solely responsible for geo-targeting.

In addition to geo-targeting specificity and inclusion of high-information maps, other WEA limitations and enhancements mentioned by the previous studies and reports include:

- **Message length:** Both NRC reports and the 2013 DHS recommendations point to the 90-character limitation of current WEA implementations as a hindrance with the proposition that a modest increase in message length might be beneficial. The 2013 NRC and DHS reports suggest the message length could effectively be extended by pagination strategies based on broadcasting multiple constituents in successive bursts. SEI's integration considerations study links the character limitation, compounded by geo-targeting imprecision, to undesirable post-emergency impact, specifically to the possibility of triggering increased voice or Internet traffic leading to congestion. A report prepared by the University of Maryland's National Consortium for the Study of Terrorism and Responses to Terrorism (UMD START) (Hamilton et al. 2014) stresses the importance of longer messages: the results of their studies indicate that longer messages "produce optimized interpretation, personalization and milling outcomes, and would likely yield maximized public protective action-taking behavior." Further, short messages were substantially less effective than longer messages at "helping people overcome their pre-conceived perceptions about different hazards and likely would be less effective at guiding people to take protective actions appropriate to the risk they face in an actual event." The Federal Communications Commission's Communications Security, Reliability and Interoperability Council is already considering an increase in the character limit.
- **Lack of understanding of AO requirements and importance of public education:** SEI's study on WEA integration considerations (SEI 2014) stresses the importance of eliciting and specifying emergency management agencies' requirements for WEA services. The UMD START report (Hamilton et al. 2014) concludes that outreach and education about the WEA service may "help to speed the rate at which members of the general public read and respond to WEA messages." Education and outreach for both AOs and the public was a main theme in our AORS as well.
- **Role of context and context-aware capabilities:** Context is any piece of information that can be used to characterize the situation of a subject of interest at a given time. It can be broadly described by the four *W*s, who, where, what and when: Identity information (who), location information (where), activity information (what) and time information (when) (Abowd et al. 1999). A context-aware application consumes the relevant pieces of context information, both static and real-time, to make a decision on behalf of the subject of interest. An example of a context-aware application is a mobile conference assistant that advises a conference attendee on which sessions to attend, where to be at a given time and with whom to connect based on the

attendee's interests, schedule and location, as well as time of the day, the published conference agenda, specific session information and presenters' and other attendees' profiles. The 2011 edition of the NRC report (NRC 2011) points to growing smart phone capabilities in terms of inferring the physical location and circumstances of the recipients and the possible role such capabilities might play in relevance-targeting of the alert messages. Examples of such context-aware capabilities for WEA include augmenting geo-targeting with client-side filtering based on the local location history of the user and the prediction of user's future location, both of which we implement and evaluate in the PUT phase of our work.

- **Interest targeting**: Currently, a WEA message can only be received in and around the affected area related to that message — leading to the *location-proxy fallacy* identified in Section 1.1. The 2013 DHS recommendations (DHS 2013) and the 2011 NRC report (NRC 2011) suggest enhancing WEA delivery by allowing the public to be notified when a WEA message is issued to their home area or to a recipient-specified area of interest. Location-history based filtering described above is also a type of interest targeting in which the recipient's geographical areas of interest are inferred rather than explicitly specified.
- **Multimedia support and presentation:** The WEA service currently supports only text messages. The 2013 DHS report (DHS 2013) recommends WEA also support richer media content in alerts, including both maps and other rich media forms such as images and audio. Inclusion of such artifacts could convey more information to the public about the situation and the required action. The UMD START report (Hamilton et al. 2014) concludes visual stimuli including color, size, shape, bolding, iconography, sound and the character of audible tones indicating the arrival of a message might influence WEA message interpretation and subsequent message response, but their effects are not yet known. We investigate the effects of such enhancements in our study.
- **Inclusion of external links:** Augmentation of emergency alert messages with external links, including URLs and social media hash-tags has been contemplated since 2009. Hughes and Palen (2009) reported on the use of Twitter in mass convergence and emergency events and found Twitter to be an important source of information broadcasting and brokerage. Their study also revealed increased use of URLs during emergency situations. The UMD START group (Hamilton et al. 2014) also studied the effects of inclusion of a URL in alert messages during a community event. Around 65 percent of users who received an alert with a link followed the link: this behavior in turn statistically significantly reduced the delay to search for additional information (milling behavior). The result was strengthened by responses from focus group participants and interviewees who predominantly indicated a preference for inclusion of URLs in alert messages. The authors concluded that consideration should be given to embedding URLs in WEA messages regardless of message length. Earlier research on accessible alerting (Mitchell, Johnson and LaForce 2010) found that users with sensory disabilities prefer to have access to a second tier of more detailed emergency information that could be provided through a URL. So an important question arises: should emergency alerts embed a URL for easy access to additional resources? We investigated this question in this study.

3. THE ALERT ORIGINATOR REQUIREMENTS STUDY

We use the term AO to refer to emergency services personnel who work in a role that involves assessing or managing emergency situations and crafting, approving and ultimately disseminating public alert messages. As a key part of our research, and in preparation for the PUT phase, the research team conducted a comprehensive study to update our understanding of the AO perspective given that WEA is currently being used in various locations. In particular, we wanted to validate certain widely held assumptions and to uncover new AO requirements.

The AORS consisted of three main components: a refreshed literature review of published information, in-depth interviews with representative AOs and an online survey of a larger community of AOs. In addition, as follow up to an open question, we investigated correlations between WEA messages and social media using Twitter data.

The details on the literature survey and the interviews were reported in a previous document titled "Alert Originator Requirements Interview Study Final Report" (Erdogmus, Griss, and Iannucci 2014). We summarize the key observations here, with more information included in the appendices. Appendix A1 provides the detailed goals, design and findings of the interviews. Appendix A2 lists the interview questions. Appendix A3 provides the questions and answers from 88 AOs across the United States who responded to our online survey.

Key observations from all three parts of the AORS pertain to the low WEA adoption rate among AOs, a resounding need for better education, a strong desire by AOs for the ability to test and evaluate system effectiveness, the WEA service's wide reach as a major advantage, the advantages of geo-targeting specificity, and strategies for increasing both the capacity of an alert to prompt action and likelihood of an alert to be relevant to the recipient via longer messages, maps and rich media.

It was surprising to learn how few AOs had actually used WEA. While a small number were approved to issue messages or were in the process of approval, many had not yet had the need or did not have the ability to issue WEA messages. Some felt that the current WEA service did not meet their needs, primarily for lack of sufficiently fine-grained geo-targeting.

Several AOs pointed out the need for better education of AOs and the public to improve the understanding, acceptance and use of the WEA service. In particular, they called for more publicity on what WEA is and under what circumstances WEA should be used instead of, or in addition to, other alerting mechanisms.

Closely tied to the publicity and education needs was the strong desire for AOs to have the ability to send test messages and receive feedback. Suggestions included conducting regular tests of the WEA service similar to the periodic tests of the Emergency Alert System (EAS). AOs believed that regular testing, with a feedback mechanism for evaluating the reach and effectiveness of the service, would increase familiarity and confidence in the WEA service.

Most AOs reaffirmed they use or plan to use the WEA service because it is able to nearly instantaneously reach anyone in a designated area (similar to "sounding a siren") without the need for prior registration. The WEA service's ability to leverage the increasing prevalence of smart phones was touted as a major advantage. AOs in general, especially those from city- and county-level jurisdictions, thought that increasing geo-targeting resolution to permit more precisely targeted alerts and alerts that could be targeted to a small area would significantly increase the effectiveness of the WEA service.

There were a variety of suggestions on improving the actionability and relevance of WEA messages. The most common suggestion was to allow longer messages: AOs in general found the current 90-character limit to be too restrictive. Some AOs also thought that it was important to permit embedding maps, web links, social media tags and other forms of rich media in alerts.

Table 3.1. Profile of AORS Interview Participants

Interviewee #	Scope	Type of Alert, Emergency or Event Handled by Interviewee Organization	Base Region	State
1	County	Shootings, Fires, Earthquakes	West	CA
2	City	Shootings, Fires, Earthquakes	West	CA
3	National	Tsunami, Hurricanes, Tornadoes, Wildfire	East	N/A
4	County	Tornadoes Hurricanes	Center	TX
5	County	Floods, Earthquakes, Tornadoes, Hazardous Materials	Center	KA
6	State	Hurricane	East	FL
7	State	Bombings, Plane Crashing, Fires, Hurricanes	East	MA
8	National	Child Abductions	East	N/A
9	Local	Fires	West	CA
10	Local	Security, Hazardous Materials, Fires, Earthquakes	West	CA
11	National	Any	East	N/A
12	Local	Security, Hazardous Materials, Fires, Earthquakes	West	CA
13	Local	Security, Hazardous Materials, Fires, Earthquakes	West	CA

Entries in roman face represent the eight AOs interviewed using the semi-structured approach.
Entries in italics represent AOs interviewed during the five additional open-format interviews.

3.1. Interview Methodology and Insights

In the first stage of the AORS, to understand the challenges underlying WEA adoption and acceptance, we studied a small group of AOs who had direct experience working with emergency alert systems, including WEA-based instances, and had experienced first-hand both the benefits and challenges of these systems. These subjects were selected from the sample population of about 600 AOs assembled from a combination of sources. They represented a mix of employees from city, county, state and national organizations. The participants' titles spanned Senior Coordinator of the Office of Emergency Services (OES), Director of OES, Program Lead of Emergency Communications, Emergency Management Liaison, Assistant Director of Emergency Management, Preparedness Coordinator, Director of National Center, OES Coordinator, Emergency Management Specialist, Emergency Operations Center (EOC) Coordinator, and Head of Protective Services. Their affiliations spanned Santa Clara County, California, City of Palo Alto, California, National Weather Service, Harris County, Texas, Johnson County, Kansas, State of Florida, Commonwealth of Massachusetts, National Center for Missing and Exploited Children, the National Aeronautics and Space Administration (NASA) Ames Research Park, and Department of Homeland Security/FEMA. Table 3.1 summarizes the demographics of the AOs interviewed.

Using qualitative, semi-structured interviewing techniques and a pre-defined set of questions, we spoke in detail with eight AOs to understand their impressions and challenges using the WEA service, document their experience of how the WEA service has been received by the general population and gather their suggestions on how the system could be improved. We augmented these interviews with data collected through five additional, informal interviews that followed an open interview approach.

Initially the purpose of this second group of interviews was to: (a) collect general information about WEA for determining the focus of semi-structured interviews and formulating formal interview questions; (b) identify further contacts who would lead us to credible subjects to interview; and (c) explore collaboration opportunities for future field trials of selected WEA improvements. Nevertheless, these interviews also provided valuable insights that complement and support those obtained through the semi-structured interviews. Therefore we included them in our analysis. Figure 3.1 shows images from the data synthesis session following the interviews.

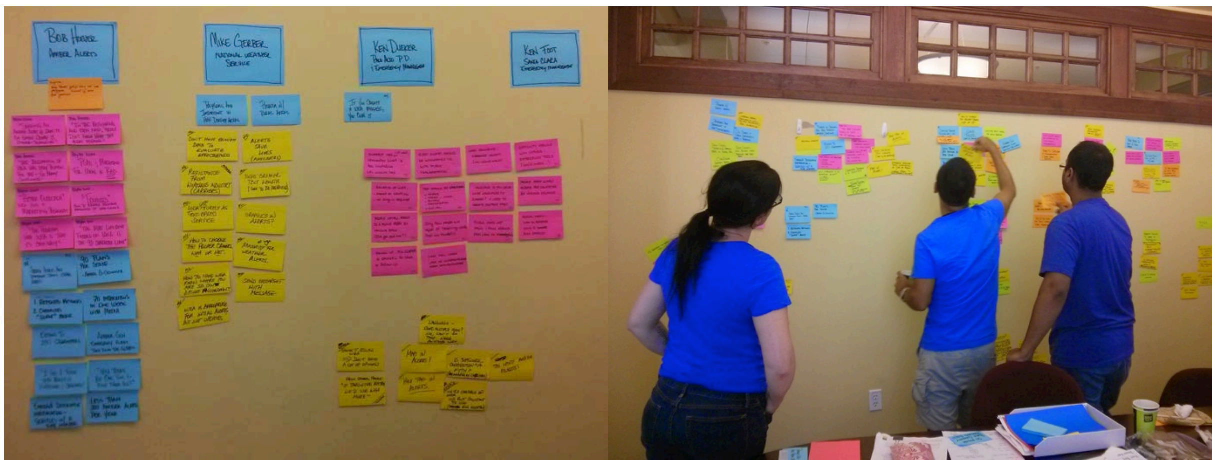


Figure 3.1. Images from the Data Analysis Session

We analyzed the collected data using the Grounded Theory method (Strauss and Corbin 1998), a qualitative approach commonly used in social sciences. The significant results were grouped into five insights that inform the future evolution of the WEA service. These insights were:

1. **Ninety characters are not enough to convey meaningful information to the public.** Most interviewees stated that they were unable to craft meaningful messages to the general population within the constraint of 90 characters. They felt allowing longer messages would be a positive step towards minimizing public confusion and increasing message relevance. This insight reconfirms a yet-unfulfilled need that the alerting community has been aware of since the inception of the WEA service.
2. **Better geo-targeting of WEA messages is a primary goal.** A majority of the interviewees stated that increased geographic precision is required in the WEA service to deliver alert messages only to those impacted by an emergency. They believe utilizing more precise geo-targeting will rectify many of the key adoption challenges faced by the WEA service.
3. **The WEA service needs to interface with social media to be relevant.** Social media was emphasized as a common tool already being used by AOs. Understanding how to align the WEA service within the constellation of existing social networks emerged as an important theme. No specific solutions were provided, however.

4. **There are two distinct conceptual models of WEA: a mere warning alarm and a richer media application with follow-through.** Questions around how to improve and evolve the WEA service uncovered two distinct mental models for the system. Some participants perceived a WEA message as a "bell ringer" technology, akin to sounding the first alarm, which relies on the public using other, mainstream communication channels to obtain additional information. Others believed that the natural evolution of the WEA service should involve uncovering ways to directly embed or reference additional information and media within the alert messages themselves and to augment them with effective incident follow-up and closure mechanisms to improve the situational awareness of the recipients.
5. **Better outreach and education for both the public and potential AOs will improve acceptance and adoption of the WEA service.** AOs agreed that not enough has been done to educate the general population around what WEA messages are, why they are important and how the public should respond when they receive a message. They also emphasized the need for better education of the AOs themselves on the benefits and uses of the WEA service. Most participants saw education and outreach essential to the success of the WEA service.

In Appendix A1, these five primary insights are accompanied by a set of implications and recommendations. In addition, there are seven secondary insights, not as strongly articulated as the primary insights. The secondary insights addressed the need to support multiple languages, technical operational issues, performance and effectiveness measurement and testing, leveraging modern smart phone capabilities with context-aware features, and ramp-up problems and growing pains in WEA adoption. Some of these secondary insights are suggestive of future areas of exploration, and were supported by answers to the online survey.

3.2. Online Survey Methodology and Insights

To gather additional data and to validate and sharpen the results from the interviews, we constructed a 30-question survey to be deployed to the larger AO community. Initial survey questions focused on demographics, WEA adoption rate and reasons for adopting or failing to adopt WEA. The next set of questions probed into general suggestions for improving WEA adoption. The remaining questions focused on the specifics of these suggestions, features desired in alert generation tools, mobile devices, and the WEA service itself, and the role of WEA within the larger landscape of emergency information dissemination.

The majority of the survey questions were multiple choice. The remaining few questions had free-text responses, requiring subsequent manual thematic analysis (Braun and Clarke 2006). We used the University of Pittsburgh's Coding Analysis Toolkit (CAT 2015) to process the free-text responses, cluster them and map the clusters to sets of emerging themes that capture their essence. This manual coding process of the free text responses reads each response and identifies words or phrases that seemed to indicate prevalent themes, iterating several times over the responses to develop meaningful clusters. For example, in one of the questions, we discovered themes that received multiple responses such as "smaller or more precise polygon," "opt out or customize message types received," "keep it short more as alarm bell," "provide enough information to cause action," "educate message creators," "educate message receivers or proactive outreach," and "do not send annoying, fatiguing irrelevant messages." Similar themes were discovered in the other free text responses.

The online survey was mailed to a list of 455 potential respondents. We received 88 responses, out of which 79 were usefully complete for a response rate of about 17%. The full set of questions and responses are supplied as Appendix A3. We summarize the main findings below, starting with the general observations and trends and continuing with specific trends that were most prominent in the responses.

Figure 3.2 illustrates two of the charts obtained from the survey quantitative analysis, with the remainder shown in Appendix A3. Survey findings are summarized below.

Respondent demographics. The majority of the respondents had significant experience in emergency alert origination, but nearly half (44 percent) did not have prior experience with WEA. Seventy percent of respondents had more than 10 years of experience in emergency services. Seventy-six percent indicated their organizations regularly issued emergency alerts. The sample was diverse, with representation from several different levels of organization: 18 percent at city level, 41 percent at county level, 33 percent at state level and 5 percent national.

Reasons for adoption and adoption failure. For the 56 percent of the respondents who used WEA, the top three reasons for WEA adoption were: (1) its ability to reach people in designated areas without prior registration (48 percent); (2) its use being mandated by other organizations (18 percent); and (3) its ability to leverage increasing numbers of WEA-capable mobile devices (16 percent). For the remaining 44 percent who did not use WEA (including those who had access to WEA, but avoided its use), top reasons for non-use included not being setup to issue WEA messages, use of an alternative service, not having training in issuing WEA messages, finding alternative services better and finding WEA geo-targeting not precise enough.

Ways to improve adoption of WEA. Fifty-one respondents provided suggestions for improving WEA adoption by AOs. The top five suggestions were to: (1) permit smaller geo-targets (61 percent); (2) better educate AOs (14 percent); (3) increase allowable message length (12 percent); (4) educate public about WEA (10 percent); and (5) provide a mechanism to test WEA messages (10 percent). Unsurprisingly, the suggestion to permit smaller geo-targets was especially prevalent among respondents affiliated with county- and city-level organizations. Fifty-four people provided suggestions for improving WEA adoption by the public. The top suggestions were to: (1) educate AOs and/or the public (24 percent); (2) allow recipients to customize which messages they receive (13 percent); (3) avoid over-warning with too many alert messages (13 percent); (4) ensure that WEA messages are actionable (11 percent); (5) improve geo-targeting (9 percent); and (6) allow longer messages (9 percent). The suggestion to avoid over-warning with too many messages correlated with the suggestion to improve geo-targeting.

Importance of geo-targeting. The ability to easily define a geo-target during alert creation was the top-preferred front-end feature for the majority of AOs, followed by compliance with the CAP. With respect to desired geo-targeting precision, 81 percent of the respondents were very likely or likely to use WEA if minimum geo-targeting resolution were less than 10 city blocks (roughly the size of an urban neighborhood), but only 61 percent were likely or very likely to use it if the minimum geo-targeting resolution were as large as 10 square miles. Sixty percent were unlikely or very unlikely to use WEA if the minimum geo-targeting resolution were as large as 100 square miles. These results indicate that AOs, in general, desire the ability to target alerts to geographical areas much smaller than what is allowed by the current WEA service. This desire was strongest among city-level respondents and to a lesser degree among county- and state- respondents.

Question 8: How likely are you to adopt and use WEA if the minimum geographic area that you can define for the delivery of the alert is:

Number of Answers: 67 (76%), Number of Skips: 21 (24%)

	Very Unlikely	**Unlikely**	**Likely**	**Very Likely**
as small as 1 square city block	11	3	15	**37**
as small as 1 neighborhood (10 city blocks by 10 city blocks)	8	4	23	**31**
1 square mile	6	5	**27**	26
as large as 10 square miles	11	10	**23**	19
as large as 100 square miles	**27**	13	9	17

Question 13: Recognizing the trade-off between message length and limits and cost of wireless broadcast technology, how effective do you believe the WEA service can be if the maximum message length is?

Number of Answers: 67 (76%), Number of Skips: 21 (24%)

	Very Ineffective	**Ineffective**	**Effective**	**Very Effective**
90 characters	3	21	**36**	5
280 characters	0	5	27	**35**
500 characters	5	16	18	**27**
1000 characters	**20**	18	12	16

Figure 3.2. Sample Survey Questions and Results

Filtering based on recipient context or message type or content. There was qualified support for giving the public the flexibility to control which alert messages they would receive. Sixty-eight percent of the respondents thought if the ability of alert recipients to filter alerts based on alert type or their own context increased, willingness to adopt and use WEA would increase. However, 32 percent of the AOs expressed significant concern if recipients could opt out of receiving some alerts by controlling which alerts they would receive or suppress in this way. City-level respondents, and to a lesser degree county-level respondents, were more in favor of some ability to customize receipt of messages based on type, content or context than were state-level and national respondents.

Message length. Although a small majority, 61 percent, of respondents felt the current 90-character maximum length is sufficient, a larger majority felt that longer messages of up to 500 characters (93 percent for up to 280 characters and 67 percent for up to 500 characters) would be more effective *even if doing so would increase the cost and push the limits of wireless broadcast technology*. This result suggests an unqualified support for increased message length in future WEA service. Messages with a 280-character limit represents the strongest preference in terms of optimal length with52 percent of respondents believing this limit would be very effective. Respondents also believed that messages of length greater than 500 characters would be ineffective.

Role and proper use of WEA. A great majority, 75 percent, of the respondents thought that the AO community must rethink the vision of WEA within the larger alert messaging landscape. Respondents were divided between whether a WEA message should simply be an alarm bell or serve a more advanced purpose. But when probed further, 99 percent agreed that a WEA message should give the recipient enough actionable information beyond simply sounding the alarm. Eighty-four percent thought that a WEA message about an impending emergency should be followed up with regular status updates, including a closure message. Also AOs agreed that the more feedback the WEA service provides about delivery of alerts, the more it would be likely to be used and adopted.

Education. Education of AOs and the public received overwhelming support from the respondents. Ninety-six percent agreed that public education on WEA is important and 92 percent thought that it had been insufficient. Similarly, 96 percent agreed that education of AOs on WEA was important and 91 percent thought that it had been insufficient.

Adding links and richer media to alert messages. AOs did not give unqualified support for enhancing WEA messages with external links and rich media. A majority of the respondents disagreed that WEA messages should be enriched with maps or images *if doing so would complicate alert generation or jeopardize the willingness of carriers to participate in the WEA service.* The respondents were divided (54 percent for vs. 46 percent against) between whether embedding links to websites and social media into alert messages would be worthwhile *if doing so has the risk of overloading the communication network*. There was strong support for integration with social media, however: 91 percent agreed that integrating WEA and social media would enhance the alerting process.

Inability to test the system. An important concern that recurred in many of the responses was the inability to test the system is a problem for several reasons, such as training message senders, increasing their confidence in use of the system and familiarizing the public. They suggested that a monthly test with the general public would be important to improved adoption. Such regular testing of the system is important both to confirm system readiness and to familiarize the public. Even today, regular tests of siren systems are performed in many jurisdictions. It was viewed as important that regular weekly (random) tests of the older Emergency Broadcast System (EBS) were performed, and since 1997 weekly and monthly regional and national tests of the newer EAS are also performed. Given the motivation and outcomes of these siren, EBS and EAS tests, WEA should also be tested regularly in a similar manner for readiness and educational benefits. Some AOs suggested the creation of a (local) testbed with a selected subset of recipients.

As might be expected, responses to certain questions were indicative of different trends based on the affiliation of the respondents; however, segmented numbers were not large enough to provide definitive analyses. These differences might be worthy of further study. For example, we saw hints of the following (unsurprising) segmentation, though the number of responses was rather small:

- City respondents, and to a lesser degree county respondents, are more in favor of some ability to customize receipt of messages based on type, content or context than are state respondents.
- City respondents, and to a lesser degree county and state respondents, favor smaller targeted areas.

3.3. Social Media Analysis

To further explore the interview finding regarding the importance integrating WEA with social media, we performed a preliminary study of the potential connection between WEA messages and Twitter posts using data from the NWS.

We analyzed Twitter social media posts that contained weather-related words such as “storm,” “rain” and “cloud,” and originated from the same areas as the NWS WEA alerts that were issued around the same time as the posts. We used real-time data collection from Twitter since it allows access to the largest volume of user-generated data for analysis. Posts were collected nationwide in real time during a two week period that spanned five NWS WEA alerts issued at different locations. For each alert, we selected posts that were created up to 24 hours before the alert was issued and up to 24 hours after the alert had expired, and that originated from within 200 kilometers of the target area of the alert. The posts then were subjected to automatic linguistic analysis to discover whether they form distinct clusters by the word usage. This allowed us to determine if there was a change in the pattern of weather-related posts before and after each alert.

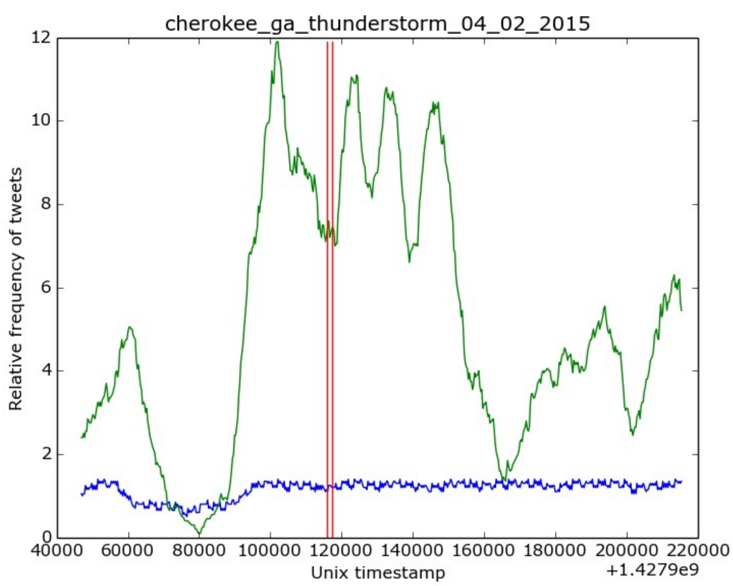


Figure 3.3. (a) Change in Twitter Posts Near a WEA Message

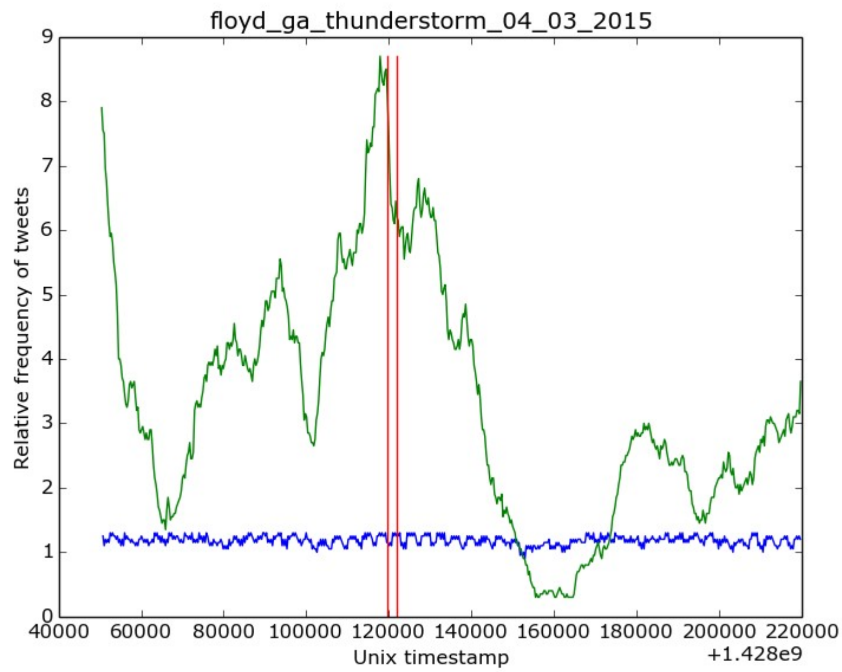


Figure 3.3. (b) Change in Twitter Posts Near a WEA Message

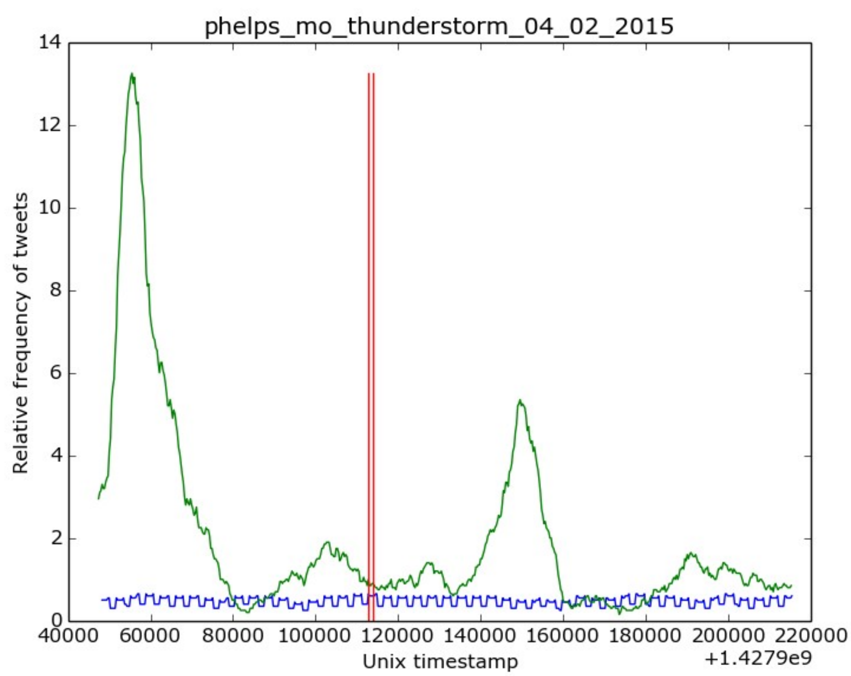


Figure 3.3. (c) Change in Twitter Posts Near a WEA Message

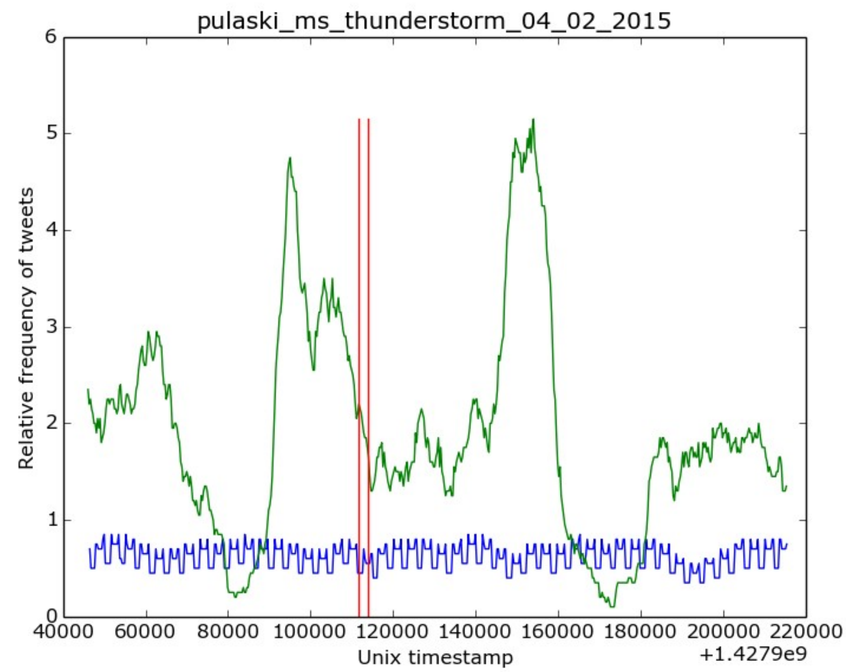


Figure 3.3. (d) Change in Twitter Posts Near a WEA Message

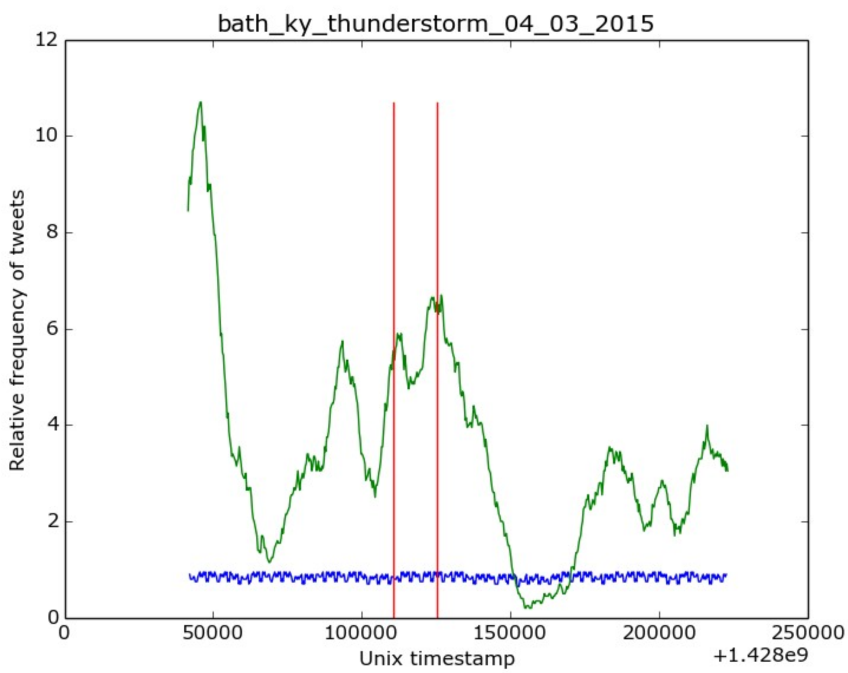


Figure 3.3. (e) Change in Twitter Posts Near a WEA Message

Figure 3.3 (a)-(e) depict the change of relative frequency of weather-related posts (i.e., the number of posts per unit of time) over time in the five different target areas. The vertical axis is relative frequency and the horizontal axis is time in seconds. All five charts depict approximately 48 hours of data. Vertical red lines delimit the times at which an NWS alert was issued and expired. Blue and green lines represent relative frequencies of posts of two different clusters that were detected based on word usage. In each chart, the relatively stable blue line most likely corresponds to posts created by an automatic system (a robot), while the green line corresponds to the posts by human users. Most charts exhibit the same general pattern of the relative post frequency by human users growing towards the middle of the chart, and then decaying. This pattern can be attributed to posters tweeting about weather more as an extreme weather phenomenon approaches, and then NWS issuing an alert near the peak intensity. There is no statistically significant change in these patterns before and after the alert, however. Counting the total number of posts before and after the alert reveals that in most cases these numbers are within one standard deviation of each other.

This small experiment inspired by an AORS finding suggests that although the public reacts to weather events on social media, there is no evidence that this response is in any way altered or caused by associated WEA alerts. In our small sample, WEA alerts failed to attract more attention to a worsening weather phenomenon than was already present. The social media buzz appears to be more immediate and the WEA alerts lagging. Our sample is too small to draw definitive conclusions, but it may be indicative of poor WEA effectiveness as compared with social media. Perhaps people are not paying as much attention to WEA alerts as they are to social media, or the service is slow to warn the public in a timely way.

4. Public Usability Trials

The purpose of the PUT phase was to test selected WEA enhancement features with human subjects. Figure 4.1 illustrates the timeline of the PUT phase.

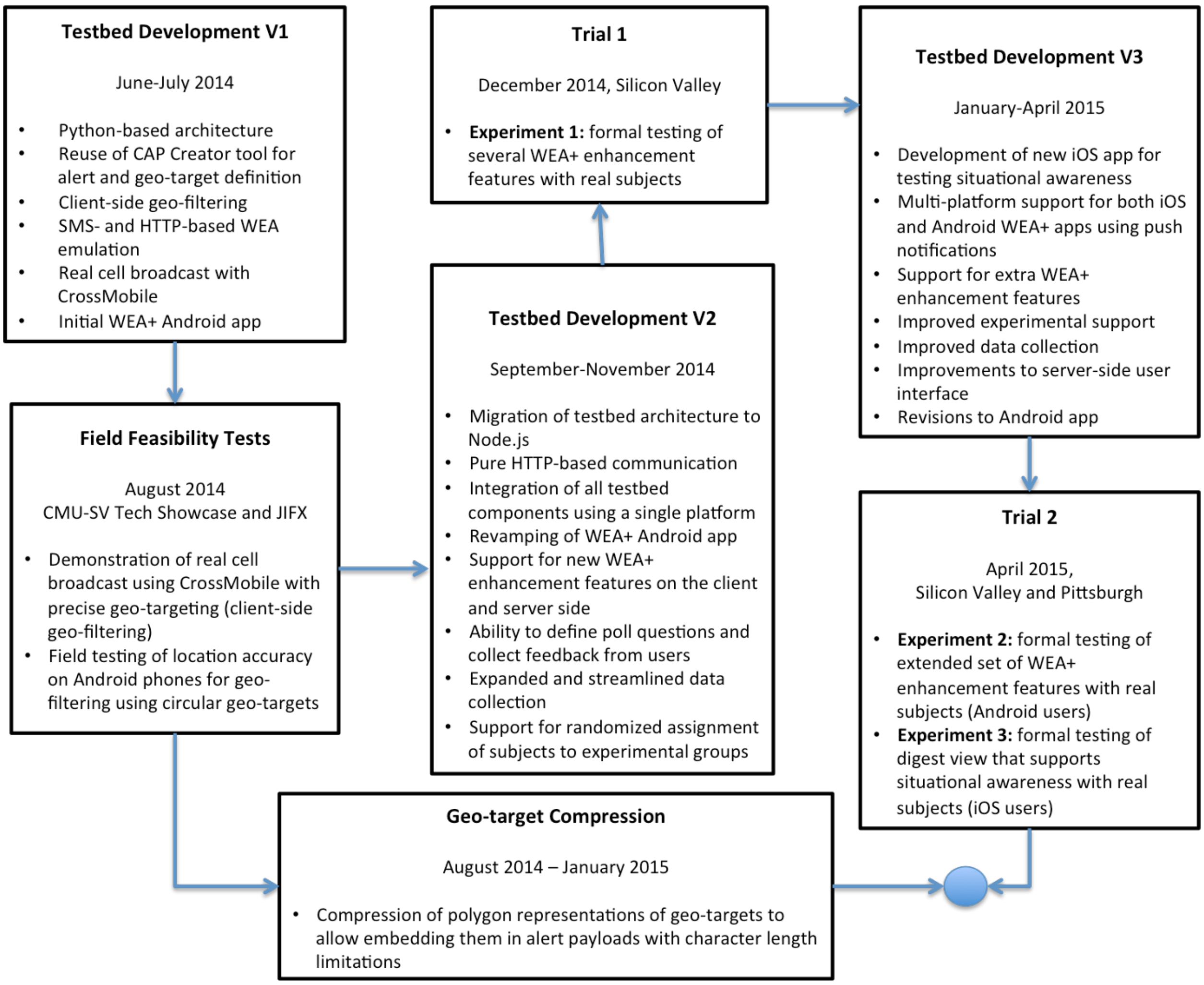


Figure 4.1. Timeline of PUT

An important activity of the PUT phase was the development of the testbed, the system infrastructure required in the two trials for delivering emulated WEA alerts to subject phones. The testbed evolved incrementally to support new trial management and WEA enhancement features in several iterations. The first version was a simple Android app and a collection of loosely coupled server-side components implemented using existing in-house and third-party tools and services, as well as glue code that connected them. Field feasibility tests demonstrated the use and accuracy of this first version with real cell broadcast technology. Successive improvements led to the second and third versions, which had an integrated, single-platform server-side architecture, and included a control center with a web-based graphical user interface. Figure 4.2 shows a screenshot of the final version of the Control Center that was used to create, schedule and manage alerts, define polygon-shaped geo-targets, and issue alerts to randomized subject groups. The third and final version of the testbed supported both Android and iOS clients in multiple geographical areas.

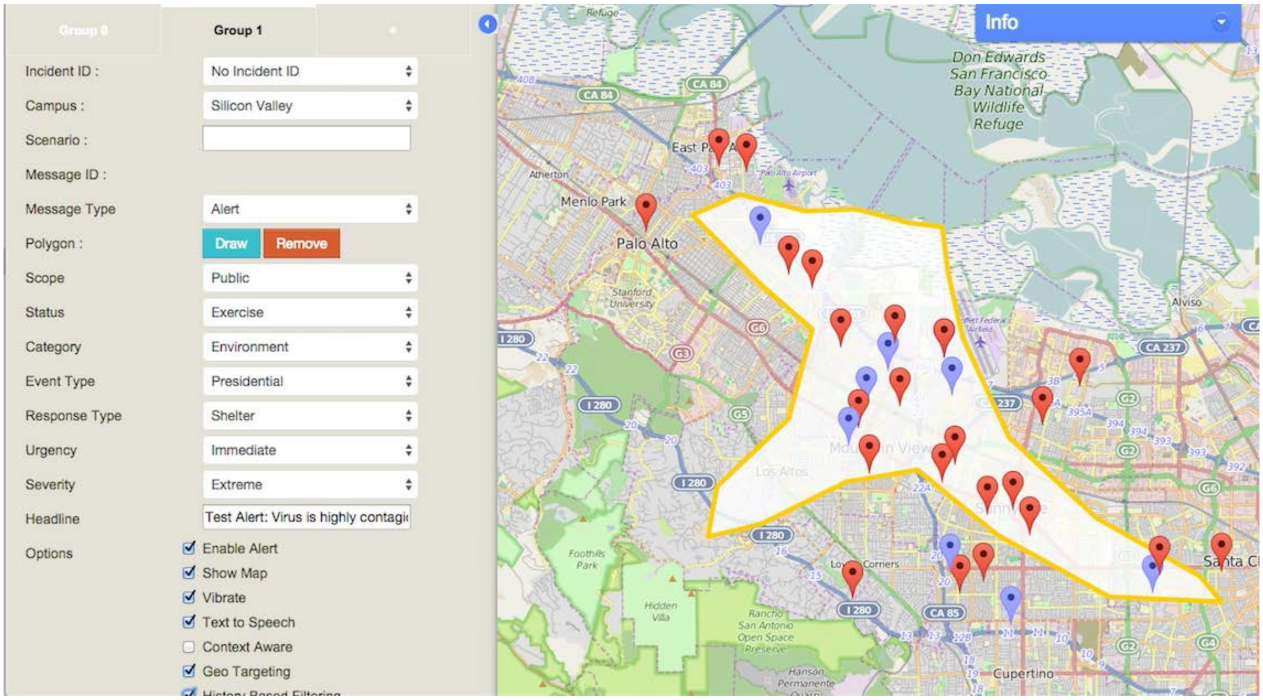


Figure 4.2. Truncated Screenshot of Control Center in Testbed Version 3 Showing the Creation of a Geo-Targeted Alert

Two trials were conducted using the second and third versions, respectively, of the testbed:

- **Trial 1** was conducted in Silicon Valley, California, with subjects recruited primarily from the Carnegie Mellon University (CMU) campus and NASA. It involved a single experiment (Experiment 1) that focused on testing features selected based on AORS preliminary findings and field feasibility tests. These features were long messages allowing more than 90 characters; precise geo-targeting with client-side filtering; inclusion of high-information maps; inclusion of external links; and alert notifications augmented with text-to-speech. Trial 1 also compared responses to initial alerts with update alerts and evaluated the effect of alert timing by comparing responses to alerts issued during the day with those issued at night. The subjects used an Android-based WEA emulation app (Android WEA+ app).
- **Trial 2** was conducted in both Silicon Valley and Pittsburgh, Pennsylvania, with an expanded sample from both locations. It involved two experiments: (a) Experiment 2, which was an extended version of Experiment 1, with subjects using Android phones; and (b) Experiment 3, with subjects using iOS phones. Experiment 3 required the development of a new iPhone WEA emulation app (iOS WEA+ app). Trial 2 focused on testing most of the features tested in Trial 1, plus a set of new features selected based on Trial 1 experience and insights, and more complete AORS findings. On Android phones, the new features supported enhanced context awareness and included the use of location history and location prediction for making better alert delivery decisions on the phone. On iOS phones, Trial 2 tested a new feature, situation digest that supports situational awareness and is meant to improve sense making in a complex emergency scenario by aggregating connected message streams.

The WEA+ apps are explained in detail in a Carnegie Mellon University-Silicon Valley (CMU-SV) technical report (Erdogmus et al. 2015). To give an idea of the additional functionality supported by the

WEA+ apps over the current built-in WEA capability of smart phones, Figure 4.3 shows a screenshot from the Android WEA+ app. It depicts an alert augmented with a high-information map: the map shows the alert's geo-targeted region as described by a polygon, the user's (phone's) location and the active period (effective and expiry date and time) of the alert. When the alert has the geo-targeting feature turned on, the app filters out the alert as long as the user is outside the geo-target and does not enter it during the alert's active period.

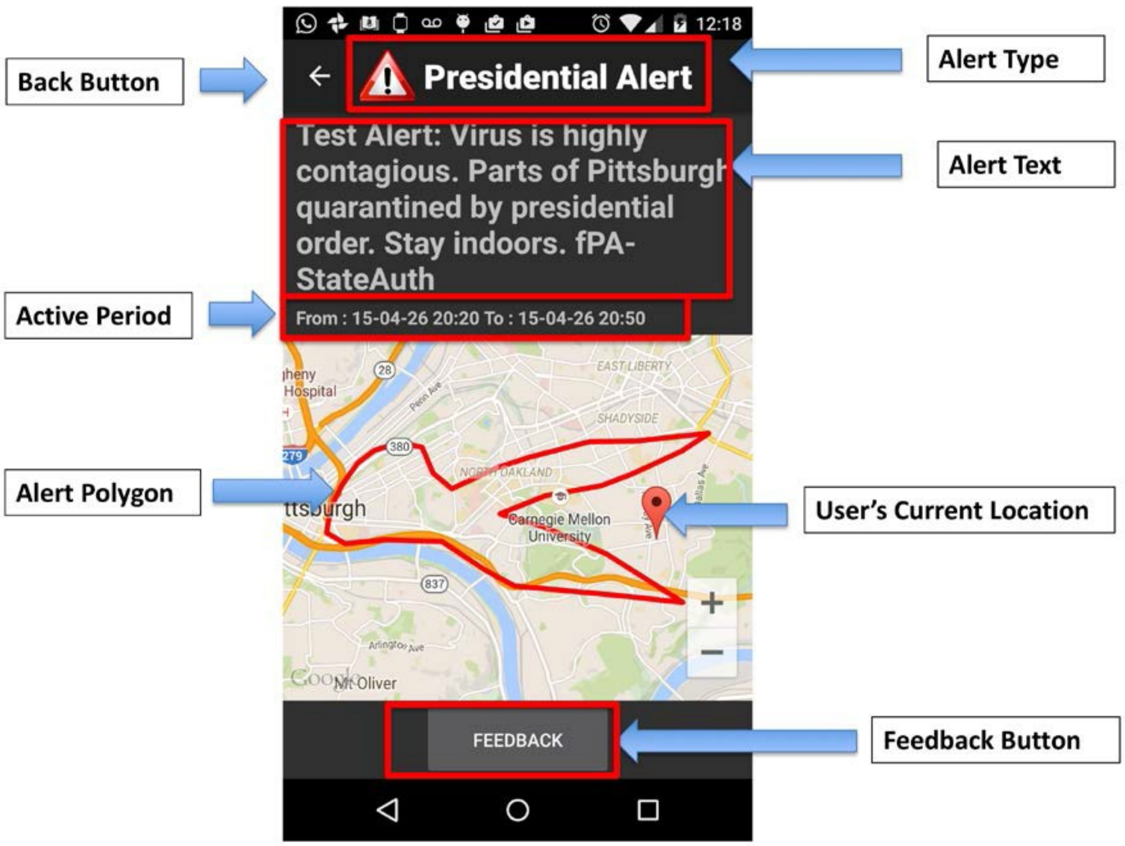


Figure 4.3. Android WEA+ App Showing an Alert Augmented with a High-Information Map

The iOS WEA+ iOS app implemented the notion of Situation Digest with a new presentation view. This feature was orthogonal to the set of features supported by the Android app. Figure 4.4 shows an example screenshot of this view. The screen is associated with a single incident in an ongoing emergency involving multiple incidents. The incident's headline is displayed at the top. The bottom panel shows the incident type and current severity in both iconic and prose forms. The middle panel shows the immediacy and the current action to be taken, again in both forms. In the digest view, the app automatically identifies the incident to which an alert update belongs and overlays the information with information from previous messages associated with that incident. The user only sees the resulting summary information — which contains all the relevant and latest information about the incident — and does not have or need access to the previous messages.

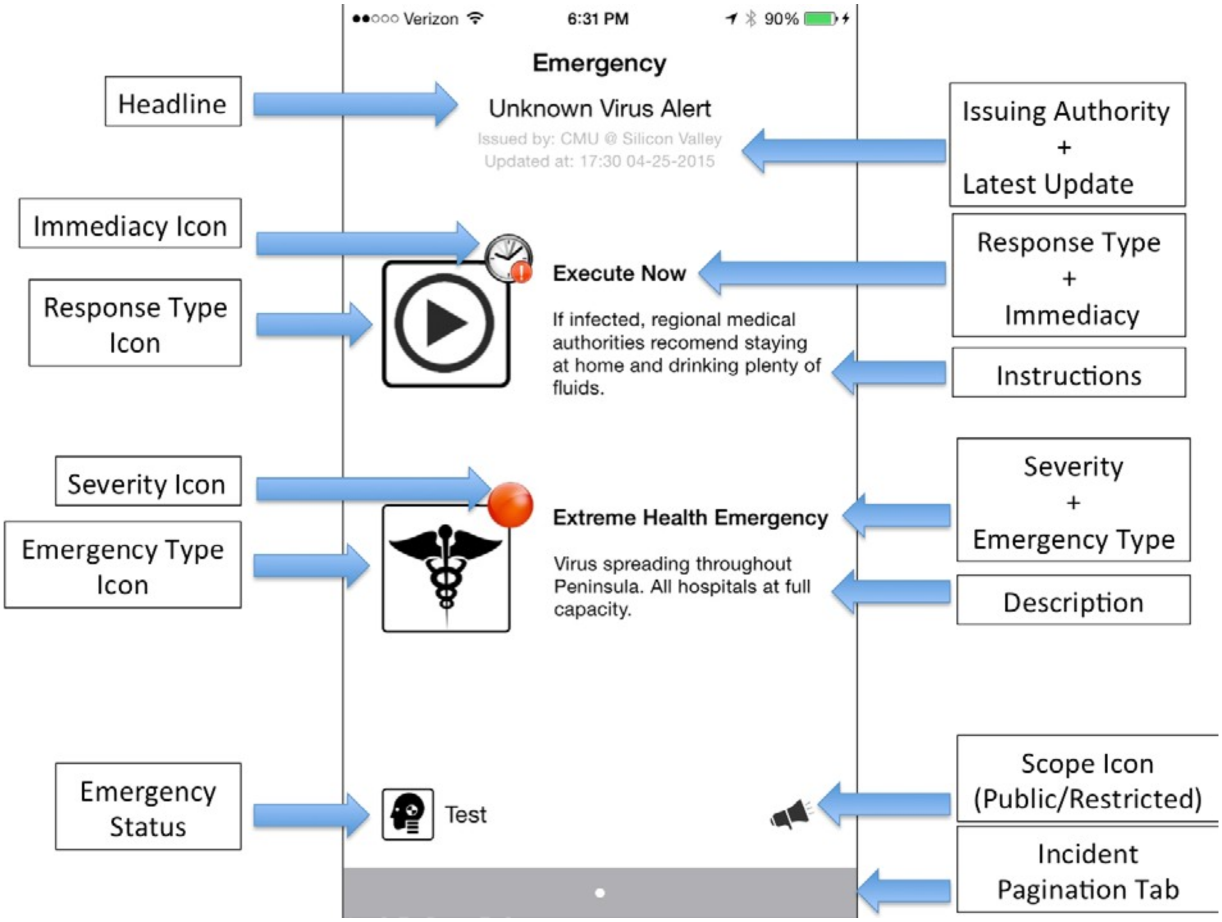


Figure 4.4. iOS WEA+ App: Implementation of Situation Digest with a New View

Both WEA+ apps also supported a data collection and feedback mechanism. Base data, such as GPS coordinates of users, timestamps and other usage statistics, were collected automatically and did not involve any explicit action from the subjects. Explicit feedback was solicited with each alert (in the Android app) or at selected points (in the iOS app) during the experiments to measure and compare certain responses. The nature of these responses is explained later in this section. Figure 4.5 shows an example feedback poll from the iOS app. The same mechanism was also used in the Android app.

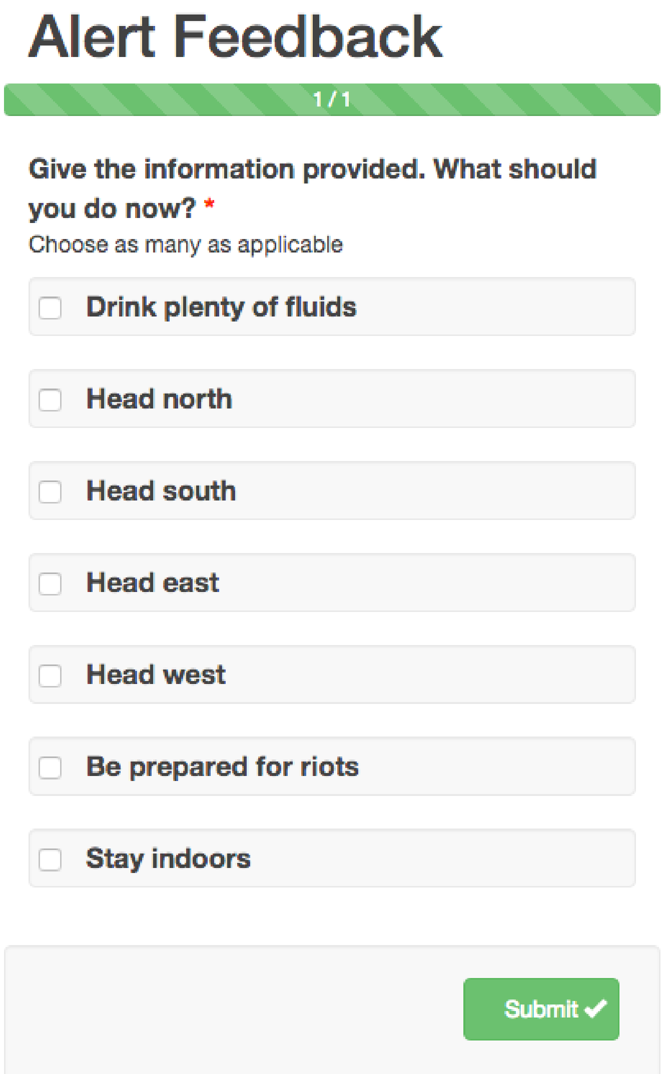


Figure 4.5. WEA+ Apps: Screenshots from Alert Feedback Questionnaire

The research questions, hypotheses, study designs and results of the experiments conducted in Trials 1 and 2 are presented in the following subsections. Details and supporting materials are included in Appendices B1 to B6.

The two trials evaluated each tested enhancement or alert characteristic, as applicable, with respect to several measured outcomes: Relevance, hindsight relevance, actionability, understanding, annoyance, adequacy, milling behavior, usefulness and situational awareness. Our top findings are:

- Fine-grained, precise geo-targeting, location-history-based filtering and the support for situational awareness markedly improve some outcomes consistently with important and demonstrable practical ramifications.
- Inclusion of high-information maps has a moderately positive effect on some outcomes, albeit not necessarily consistently, with moderate practical ramifications.
- Long messages, TTS notifications, location-prediction-based filtering and inclusion of external links have little, no significant or mixed effect on most outcomes.

A parallel side activity of the PUT phase focused on geo-target compression. A central crosscutting idea in Experiments 1 and 2 was the transmission of the geo-target with the alert message. Several context-aware enhancement features, including precise geo-targeting, high-information maps, location history and location prediction, relied on this mechanism. The geo-target compression activity explored the feasibility of embedding realistic polygon-shaped geo-targets in size-limited alerts for processing by the recipient phones to implement a variety of client-side, context-aware improvements. The results were extremely promising. Therefore, we summarize them in the last subsection.

4.1. Research Questions for Trials 1 and 2

The following high-level research questions were tackled in the two trials of the PUT phase:

- Potential Enhancements: Would certain improvements to the WEA service make it more effective for the public?
- Alert Characteristics: Do certain alert characteristics impact the effectiveness of the WEA service for the public?
- Overall Impressions: After exposure to the WEA service, does the public appreciate the benefits of the WEA service?

For the first two research questions, a set of potential enhancements (termed *enhancement features)* and *alert characteristics* were pre-selected for study based on a literature review, AORS interview and survey findings, technical feasibility and ability to leverage modern smart phone capabilities. Insights from Trial 1 also lead to new features to be covered in Trial 2. We present the enhancement features and alert characteristics studied first. Then we explain how *effectiveness* was formally evaluated in terms of a set of outcome measures.

Not all enhancement features were equal in terms of their feasibility, complexity, nature and consequences. Some features were easy to implement with very minor modifications to the current WEA service. An example of such a feature is the use of TTS in alert notification. Such a feature is straightforward to implement using client-side changes only (the built-in WEA app on smart phones). It does not require any other changes to the current WEA service, including the alert creation process and the delivery infrastructure. On the opposite end of the spectrum is Situation Digest, which is not a stand-alone feature in the same sense of TTS, but more another way of thinking about emergency alerts. Such a feature cannot simply be implemented with minor changes on the client side since it affects the alert creation process, requiring AOs to specify certain metadata required to organize and connect alerts that belong to the same ongoing emergency incident in a systematic way. It also requires special encoding, and possibly subsequent compression and decompression, to be able to bundle the metadata with the alert text for transmission through the system while obeying the existing character length limitations. This may have further implications with respect to alert authentication and authorization, which affects the normal workflow. Other tested features, such as precise geo-targeting, fall somewhere in between these two ends of the feature spectrum in terms of feasibility, complexity, nature and consequences.

4.2. Independent Variables and Hypotheses

We refer to the enhancement features and alert characteristics collectively as *tested factors*. These represent the independent variables. Each tested factor takes a binary value representing either the presence or absence of the underlying feature or characteristic.

The factors selected for testing are listed in Table 4.1 (alert characteristics) and Table 4.2 (enhancement features) together with a description, rationale and hypothesis for each factor.

Table 4.1. Alert Characteristics Studied

Tested Factor	Rationale and Assumptions	Hypothesis
Timing **The time period within which an alert is active.**	The timing of alert messages may influence recipients' reaction to an alert. Are there any differences between alerts received at different times of the day in terms of how recipients perceive them? If there are such important differences, alerts can be customized or embellished with features based on their timing to improve outcomes that are compromised by timing.	There is a difference between the effectiveness of alerts received during the day and alerts received after hours.
Update Alert **An alert that modifies the information sent in a previous alert.**	The majority of AOs believe that alerts informing citizens of an impending or current emergency should be followed up with alerts that update the status of the emergency. It may be difficult to convey the context and history of an ongoing or past emergency in a short follow-up message, however. Are alerts that are status updates as effective as initial alerts?	There is a difference between the effectiveness of initial alerts and update alerts.

Table 4.2. Enhancement Features Studied

Tested Factor	Rationale and Assumptions	Hypothesis
Long Message **Alert messages longer than 90 characters (Experiment 1) or longer than 130 characters (Experiment 2).**	Current WEA service limits alert messages to 90 characters. Future WEA service is likely to support longer messages. Most AOs find this limit too restrictive and advocate longer messages with more information content.	Longer alert messages containing more information about the underlying emergency are more effective.
High-Information Map **A map showing the recipient's location and the alert's target area.**	Current WEA service only supports text. AOs advocate rich-media support that includes pictures and maps. Previous research points to the value of high-information maps. Modern smart phones are able to cache maps and have GPS and location capabilities. Future cellular technology will provide better support for maps. Compression techniques make it feasible to embed the geo-target into the alert payload.	Alerts displaying high-information maps overlaid with the targeted geographical area of the alert and the position of the recipient are more effective.

Tested Factor	Rationale and Assumptions	Hypothesis
Geo-targeting **Fine-grained, precise targeting of an alert to a geographical area specified as a polygon. Geo-target is embedded in the alert message and delivery decision is made on the recipient's device based on device location.**	Current WEA service does not allow precise geo-targeting of alert messages to a designated alert area. AOs believe future WEA service should address this limitation. Modern smart phones have GPS and Wi-Fi location capabilities that allow on-device filtering based on recipient's location, provided that a representation of the targeted area is transmitted with the alert. Compression techniques make it feasible to embed the geo-target into the alert payload. Most users keep location services enabled on their phones, allowing the phone to determine device location.	Alerts that are precisely targeted to recipients in a specific alert area are more effective.
External Link **Inclusion in the alert message of a URL or a social media (e.g., Twitter) tag that points the recipient to an external source for further information.**	Currently inclusion of clickable links to websites and references to social media tags are disallowed or discouraged in WEA messages due to possible network congestion that may be caused by too many citizens simultaneously trying to access the same internet resources.[7] Inclusion of such external information requires additional characters. Some AOs believe that external links in alerts may improve the effectiveness the WEA service by allowing citizens to easily access information that may help them better assess the underlying emergency. Are external links worth including in alert messages?	Alerts containing external links are more effective.
Text-to-Speech (TTS) **Recipient is notified of an incoming alert via special ringtone, vibration and spoken text**.	Recipients alerted using different tonal modes might react to alert messages differently and at different speeds and rates. Is there a difference between alerts notifying the recipient using vibration and a ringtone and those notifying the recipient with vibration and the spoken alert text instead of a ringtone?	Recipients react to TTS alerts faster, and there is a difference between the effectiveness of alerts using TTS and those using vibration for notifying the recipient.

[7] Restricting links to specialized short web pages cached near cell towers would be a resource-light option, but this option is seldom considered.

Tested Factor	Rationale and Assumptions	Hypothesis
Location History **Geo-targeting takes into account recent location history of the recipient. If the recipient has recently visited the targeted area, the alert is delivered even if the recipient is outside it during the alert's period.**	Location history supplements geo-targeting by targeting those recipients who frequently visit the alert region, thereby reaching a wider audience interested in the alert. For example recipients in their offices would still be interested in a fire next to their homes, which would not be received if just geo- targeting based filtering was used, but would be received if location history based filtering is used.	Alerts targeted to recipients who frequently visit the alert area are more effective.
Location Prediction **Geo-targeting takes into account the direction of the recipient's movement. If the recipient is currently outside the targeted area, but moving towards it and is predicted to be in the alert area within the alert period, the alert is delivered.**	Geo-targeting and Location History (as discussed above) use recipients' current and past location information to filter more relevant alerts; however, they do not consider users' future location in assessing alerts relevance (e.g. a recipient moving towards a fire region could still be interested in a fire alert targeting that region, even though the recipient is not in the geo-targeted region or has not visited the geo-targeted region in the recent past). Location Prediction uses recipients' movement information (such as speed and direction) to predict their presence in the alert region in near future.	Alerts targeted to recipients who are moving towards an alert area are more effective.

Tested Factor	Rationale and Assumptions	Hypothesis
Situation Digest **Alert streams are grouped into incidents. Client device has a digest view for each incident that displays the latest status of the incident. Each incoming alert updates the digest view of the incident to which it belongs and the recipient is shown the digest view for that incident.**	WEA service may be suitable for alerting the public via individual, independent messages in short-duration, isolated emergencies. However, cognitive burden on the recipient may increase and effective sense making may become difficult when multiple updates continually modify different aspects of a complex, large-scale and evolving emergency. Consequently, recipients' situational awareness may be compromised. In these situations, such as a serious earthquake that may snowball into multiple, interrelated incidents, expecting users to mentally stitch together sequences of past alerts to assess the latest situation may lead to confusion and incorrect or inappropriate response. Can alternative alert generation and alert presentation schemes that minimally modify the current WEA service improve the recipients' situational awareness in complex, evolving	In complex emergencies, a modified WEA service that provides scenario-based digest information to the recipients increases the recipients' situational awareness as compared to independent, interleaved sequences of individual alert messages as used in the current WEA service.

4.3. Dependent (Response) Variables

In the experiments, *effectiveness* was defined in terms of a set of outcome constructs, which were selectively applied to the tested factors depending on whether they were meaningful for the tested factor and whether they were feasible to study within the constraints of the experimental setting. These outcome constructs represent the dependent response variables and are quantified by explicit feedback from the subjects in response to receiving an alert or in response to a timed poll within an emergency scenario. The components of effectiveness are:

- **Understanding**: Does the presence of the tested factor in an alert make it easier to understand for the recipient as compared to an alert in which the tested factor is absent? Does it make it less confusing?
- **Relevance**: Does the presence of the tested factor in an alert make it make it more relevant to the recipient as compared to an alert in which the tested factor is absent?
- **Hindsight relevance** (used only with geo-targeting): Was an alert targeted to a specific geographical area relevant to others who were outside the geographical area?
- **Annoyance**: Would the alert, in the way it was delivered and presented, annoy the recipient in a similar real emergency situation?
- **Actionability**: Did the recipient find the alert actionable in terms of knowing what to do next?
- **Milling behavior**: Milling refers to seeking confirmation from alternative sources when presented with a piece of information to improve sense-making (Bourque et al. 2013). Does the presence of the tested factor in an alert encourage milling behavior compared to an alert in which the tested factor is absent?
- **Adequacy**: Does the alert contain sufficient information enabling the recipient to assess it?
- **Usefulness**: Did the recipient find the alert useful?

- **Response delay**: The elapsed time between an incoming alert notification and the reaction of the recipient as recorded by opening the alert or giving feedback.
- **Response rate**: The percentage of recipients who received an alert and responded to it by opening the alert or giving feedback.
- **Situational awareness**: Does the presentation of alert messages (as a single stream or using a digest view) associated with an ongoing emergency affect the recipient's ability to make sense about the evolution of that emergency?

With respect to milling behavior, we assume that it is a positive response to an alert and is a proxy for a recipient's active interest. Kelly et al. (2010) suggest that milling behavior is positively correlated with actionability. They point out to several references in the warning literature, e.g., (Mileti and Fitzpatrick 1992; Mileti and Darlington 1997) which "has found that the more searching for information that a person does, the more likely he/she is will be to respond to a warning message" and "searching for more information about getting ready for earthquakes was positively and significantly associated with preparedness and mitigation actions." Based on the literature, we maintain this assumption. However, this assumption may be construed as somewhat paradoxical since if an alert should be self-sufficient, milling behavior should be unnecessary.

In addition to alert-based outcomes, we evaluated *general outcomes* at the end of each experiment using a post-test survey (the final questionnaire). These outcomes were related to overall impressions about WEA and preferences regarding selected tested enhancement features.

- **WEA Benefits:** After a week of exposure to frequent simulated WEA messages, the subjects were asked about their overall impressions about the benefits of the WEA service by answering a pair of questions:
    - Do you believe wireless emergency alerts are useful?
    - Do you believe wireless emergency alerts could save lives?
- **Feature Preference:** Preferences regarding selected features to which the subjects were exposed (inclusion of high-information maps, inclusion of external links, location history, location prediction and situational awareness) were assessed using the questions below. Randomization ensured that all subjects were exposed to all enhancement features and had a basis for comparison by receiving alerts without each of the tested enhancement.
    - High-Information Map: Did inclusion of maps with some alerts increase their clarity and relevance?
    - External Link: Was inclusion of links to external sources (URLs or social media tags) useful?
    - Location History: Some alerts targeted to a specific geographic area were delivered to you if you had recently visited the targeted area even if you were outside it when the alert was sent. Was this feature useful?
    - Location Prediction: Some alerts targeted to a specific geographic area were delivered to you if you were moving towards the targeted area even if you were outside it when the alert was sent. Was this feature useful?
    - Situational Awareness: You were presented with more than one view to represent the messages. One was a stream of messages and the other was an updating situation panel. Which one do you prefer?

A summary of the experiments conducted during the two trials, as well as tested factors and outcome constructs evaluated for each experiment are given in Table 4.3. An "X" in a cell indicates that the enhancement feature in the corresponding row was tested in the experiment of the corresponding column. Experiment 1 had an explicit separation between messages within the standard 90-character limit and longer messages. Experiment 2 did not have an explicit threshold and the short-long threshold was inferred from the data.

The final questionnaire also addressed study validity. The design and execution of the experiments were validated through a set of indicators, listed in Table 4.4. These indicators relate to clarity of motivation for the trials, frequency of alerts issued, intrusiveness of the experiments, realism of the alerts issued, ease of use of the WEA+ mobile apps and effectiveness of the trial kickoff event. Each indicator was measured through a corresponding question in the final questionnaire. Study validity is discussed in Section 4.13.

Table 4.3. Summary of Experiments: Experiment Locations, Mobile Platforms Used by Subjects, Distribution of Subjects, Distribution of Alerts, Features Tested, Characteristics Tested and Outcomes Evaluated

Trial	**Trial 1**	**Trial 2**			
Locations	Silicon Valley (SV)	Silicon Valley (SV) and Pittsburgh (Pgh)			
Experiments	Experiment 1	Experiment 2		Experiment 3	
Platforms	Android	Android		iOS	
# Subjects	52	SV	Pgh	SV	Pgh
		42	46	54	43
# Alerts Issued per Subject	24	54		60	
Enhancement Features Tested					
Long Message	X (Explicit)	X (Implicit)			
High-Information Map	X	X			
Geo-targeting	X	X			
External Link	X	X			
Text-to-Speech	X	X			
Location History		X			
Location Prediction		X			
Situation Digest				X	
Alert Characteristics Tested					
Timing	X	X			
Update Alert	X				
Alert- or Scenario-Based Outcomes					
Understanding	X	X		X (via Situational Awareness)	
Relevance	X	X			
Hindsight relevance	X (Geo-targeting)	X (Geo-targeting)			
Annoyance	X	X			

Trial	Trial 1	Trial 2	
Actionability	X	X	X
Milling behavior	X	X	
Adequacy	X	X	
Usefulness	X (High-Information Map and External Link)	X (High-Information Map and External Link)	
Situational awareness			X
Response delay		X (Text-to-Speech)	X
Response rate		X (Text-to-Speech)	
Outcomes Related to Overall Perceptions			
WEA Benefits	X	X	
Feature Preference	X (High-Information Map and External Link)	X (High-Information Map, External Link, Location History, Location Prediction)	X

Pgh: Pittsburgh; SV: Silicon Valley

Table 4.4. Study Validity Indicators Assessed in the Final Questionnaires.

Study Validity Indicators			
Trial	**Trial 1**	**Trial 2**	
Locations	Silicon Valley (SV)	Silicon Valley (SV) and Pittsburgh (Pgh)	
Experiments	Experiment 1	Experiment 2	Experiment 3
Clarity of Motivation	X	X	X
Frequency of Alerts	X	X	X
Intrusiveness	X	X	X
Realism	X	X	X
Ease of Use	X	X	X
Kickoff Event Effectiveness	X		

4.4. Trial 1 Preparations and Demographics

Trial 1 (Experiment 1) was conducted with volunteers from CMU's Silicon Valley campus. Subjects included students, faculty and staff. Ten days before the trial start, the research team conducted an online pre-trial survey of the campus population to assess smart phone usage and determine eligible users with Android phones. Fifty-six responses were received, of which 50 percent were Android users and 50 percent were iOS users. Since the experiment needed the subjects to use data and location services, the eligibility survey also inquired about data plans and usage patterns. Seventy-five percent of the respondents had both data and text plans and 95 percent indicated that they kept data services enabled outside Wi-Fi zones. Fifty-eight percent kept GPS/location services enabled most of the time and 34 percent indicated that they kept these services enabled some of the time. Thirty percent said they would volunteer for a research experiment with smart phones and 52 percent said they might volunteer.

After the eligibility survey, posters, on campus email lists and word of mouth officially advertised the experiment. A week before the kickoff, an online signup form was deployed to formally solicit volunteers. Forty-three positive responses were received, with two volunteers indicating they would need loaner Android phones. To increase participation, and anticipating further requests for loaner phones at the kickoff, the research team decided to prepare 25 Android phones to loan to prospective volunteers.

On the evening of day 1 of the trial, a kickoff event was held during which the research team explained the project goals and the experiment to the attendees; explained how privacy, security and confidentiality would be handled; demonstrated the WEA+ Android app; completed the sign-up procedures including signing informed consent forms; gave instructions; and helped the volunteers install the WEA+ app and set up their phones. Test alerts were sent to test the WEA+ app and allowed the participants to become familiar with the app. About 60 volunteers attended the kickoff event and 43 signed up for Trial 1 on the spot. The research team loaned 15 Android phones with SIM cards and data plans to volunteers who did not have Android phones. An additional nine volunteers who could not attend the kickoff event signed up on day 2, bringing the total number of subjects to 52 (six faculty/staff and 46 graduate students). One of these volunteers' phones failed to register with the WEA+ server so we collected data from a total of 51 phones over six days. This corresponds to a 16 percent participation rate from the sampled population of 324 students, faculty, scientists and administrative staff. Actual participation peaked at 48 during the trial based on active phones detected from day to day. Most users were active every day, but some dropped out temporarily on certain days.

The participation instructions asked the subjects to keep their phones on and data and location services enabled throughout the experiment. Subjects were also asked to charge their phones each night. Installation instructions, the informed consent form and other trial-related information were posted at the project website (weacmu.org). During the trial, the project website was updated with fresh posts to keep the participants informed and motivated.

During Trial 1, subject participation levels stayed relatively steady, peaking around the second and third day. We tracked participation starting the day after the kick-off event and stopped on the last day of the trial (day 8).

4.5. Trial 2 Preparations and Demographics

Trial 2 (Experiments 2 and 3) followed a similar preparations schedule to Trial 1, but was conducted simultaneously in two separate locations on two different smart phone platforms to test two disjoint sets of features. Experiment 2 was a partial, but larger replication of Experiment 1 and used a new version of the WEA+ Android app. Experiment 3 was new and tested a different paradigm via a new mobile WEA+ app deployed on iOS for iPhone users.

Like in Experiment 1, Silicon Valley subjects were solicited from CMU's local campus and targeted students, staff, scientists and faculty. We did not deem this to be sufficient because of repeat targeting of the same population, however. To extend the sample, we mobilized our local emergency services and NASA Ames Research Center contacts to recruit extra subjects through an email, flyer and word-of-mouth campaign. Two kickoff/information meetings were held, one at the CMU campus the evening of the trial start and another at a Community Emergency Response Team (CERT) training event in neighboring Sunnyvale the week before. We reached out to Amateur Radio Emergency Services (ARES) volunteers as well.

We changed our signup protocol to allow online signup (approval from CMU's Institutional Review Board was obtained for protocol changes). The project website was updated to include an online signup form, a complete set of written instructions, all information packages (consent information and privacy/security policy), and instructional videos for both versions of the WEA+ app. A research team member also attended a Red Cross event the week before to distribute flyers and advertise the trial.

In Pittsburgh, subjects were recruited from CMU main campus population, including the Software Engineering Institute, through an email and flyer campaign. A separate kickoff/information meeting was held on campus in Pittsburgh the evening of the trial start. One research team member was based in Pittsburgh and another traveled to Pittsburgh to help with logistics, recruitment and the kickoff event.

The additional recruitment activities, efforts and methods resulted in an increase in sample size compared to Experiment 1. A total of 185 volunteers signed up to be subjects in Trial 2, of which 96 were in Silicon Valley and 89 were in Pittsburgh. Of the 96 Silicon Valley subjects, 42 were Android users and assigned to Experiment 2 and 54 were iOS users and assigned to Experiment 3. Fifty percent were CMU students, faculty or staff and 50 percent had non-CMU affiliations (NASA, CERT volunteers, ARES volunteers, employees of various local emergency services organizations, and city employees including the cities of Mountain View, Palo Alto and Sunnyvale). Of the 89 Pittsburgh subjects, 46 were Android users and assigned to Experiment 2 and 43 were iOS users and assigned to Experiment 3. All Pittsburgh subjects had CMU affiliations. Since our Pittsburgh email campaigns specifically targeted Electrical and Computer Engineering and Computer Science populations, we estimate that most subjects belonged to these departments. In both locations, CMU subjects were predominantly students as opposed to staff or faculty. Across both locations, we were able to recruit 88 subjects for Experiment 2 and 98 for Experiment 3. Despite this higher than expected signup rate, actual participation varied from day to day, and not all users were active each day. Participation peaked at 66 for Experiment 2 and 65 for Experiment 3. The number of actual registered phones from which we were able to collect data was 72 for Experiment 2 and 98 for Experiment 3.

Since we could support two major smart phone platforms, no loaner phones were used during Trial 2.

We also collected demographics regarding gender and ethnicity during the sign-up process. Respectively, 70 percent and 66 percent of Silicon Valley and Pittsburgh subjects were male. Thirty-five percent of Silicon Valley subjects identified themselves as belonging to a visible minority group (either Asian, Black/African American or Hispanic) compared to 55 percent of Pittsburgh subjects. However, 55 percent and 39 percent of the respondents in Silicon Valley and Pittsburgh, respectively, did not indicate their ethnicity in the sign-up form. Asians represented the majority of those who identified themselves with a minority group (85 percent in Silicon Valley and 87 percent in Pittsburgh.)

Instructional content provided via the website and kickoff/information meetings to the subjects was similar to content provided in Trial 1, except for the additional instructions and demonstrations for the new iOS WEA+ app and updates regarding changes to the installation procedures. A detailed Frequently Asked Questions page was created and posted at the project website (weacmu.org). All trial-related information, including participation and installation instructions, instructional videos, forms, confidentiality, privacy and security information were posted at the website. The website was updated regularly with fresh posts to keep the subjects engaged and informed.

As in Trial 1, subject participation during Trial 2 was steady, again peaking around the second and third day of the trial. In Trial 2, we started tracking participation a day earlier than Trial 1, beginning on the evening of the kick-off event.

4.6. Design of Experiments 1 and 2

Experiments 1 and 2 were conducted with subjects using the Android version of the WEA+ mobile app. They were similar and used the same single-factor, randomized repeated-measures design. The single factor was binary, representing either the control (Group A) or a tested feature (Group B). Multiple alerts were sent over multiple days.

Figure 4.6 illustrates the design of Experiments 1 and 2.

In these experiments, each alert had two variations for A/B-style testing of differences between the variations. Subjects were evenly and randomly divided into two groups, Group A and Group B, for each alert. A standard *control* alert was sent to subjects in Group A and a *variation* alert, an enhanced alert with a predetermined tested feature, was sent to subjects in Group B. Subjects were randomly re-assigned to Group A or Group B for each alert, so the groups randomly changed from alert to alert, but the distribution remained as even as possible for each alert. Group A subjects received the control alert and Group B the enhanced alert with the tested feature. The tested feature for Group B also changed from alert to alert. This way, the subjects were randomly exposed to a series of new features over the course of each experiment, resulting in a *repeated-measures, cross-over* design. In repeated-measures, cross-over studies, each subject is randomly assigned to a sequence of treatments, which includes a set of tested treatments and a control treatment to which each of the tested treatments are compared. In our case, a tested treatment corresponds to an alert having a combination of enhancement features and the control treatment corresponds to an alert with no enhancement features. Thus each subject crosses over from one treatment to another as they receive alerts with different combinations of enhancement features or with no features.

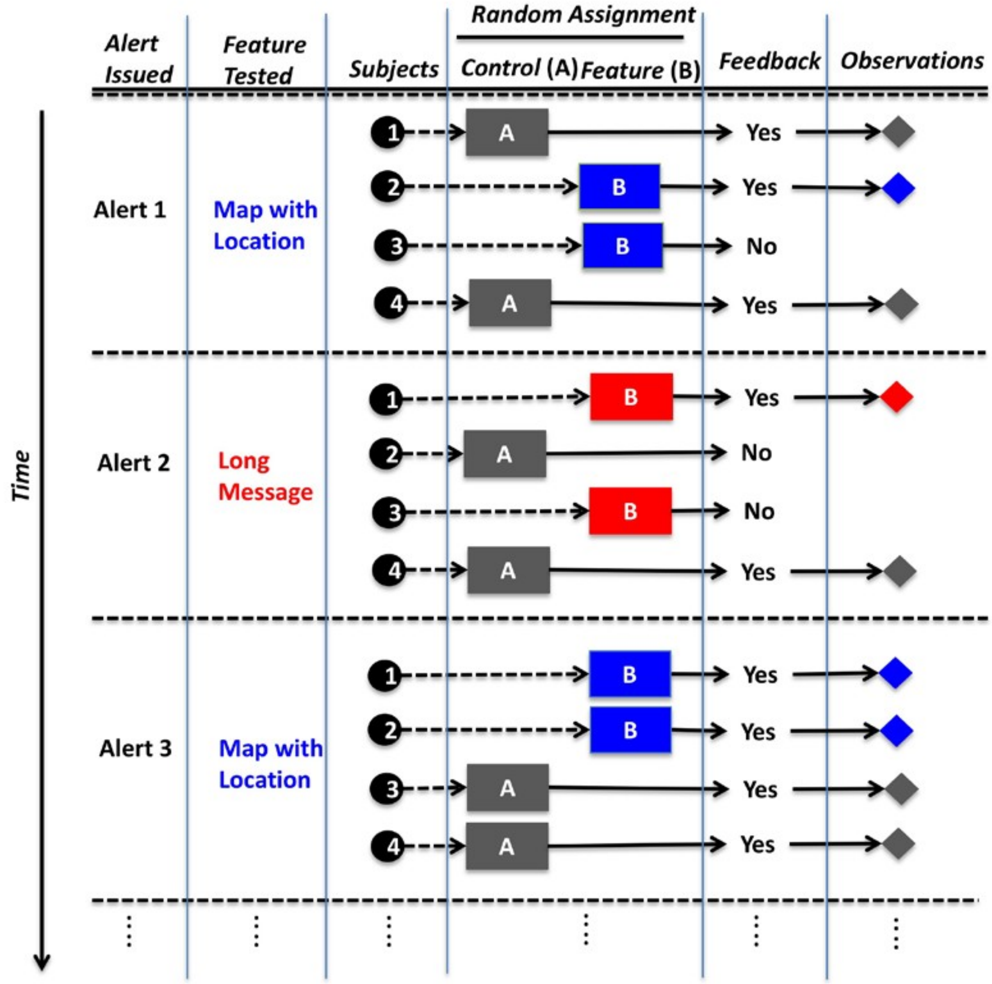


Figure 4.6. Experimental Design Used with Android Subjects in Experiments 1 And 2 to Test Alert-Based WEA Enhancement Features

In Experiment 2, alert characteristics and tested features were the same across the two locations, but alert contents were adapted to use local landmarks and geography and timed to make sure matching alerts were received around the same local time in Pittsburgh and Silicon Valley. In both locations, each feature was tested with at least three enhanced alerts.

After each alert, the WEA+ app asked the subjects to provide feedback on the last-received alert by answering a set of multiple choice questions on their phones to assess alert-based outcomes (understanding, relevance, annoyance, actionability, milling behavior, adequacy and usefulness) for the tested feature relative to the control. The feedback questions used in Experiments 1 and 2 are listed in Table 4.5. Subjects could also provide a free-text comment about each alert. If a subject chose to provide feedback on any of the evaluated outcomes, an observation was recorded for the control or the tested feature for the alert-subject pair. Observations from alert pairs (control and enhancement) testing the same feature were pooled for analysis of that feature. A control alert used in testing one tested feature was also included as control in the analysis of another tested feature only if the two features did not interact to ensure the analysis of a feature is not confounded by the presence of extra features.

Table 4.5. Alert-Based Feedback Questions for Experiments 1 And 2 to Evaluate Outcomes Enhancement Features

	Alert-based Feedback Questions	
Question	**Outcome Construct**	**Applicable Features/Groups**
Did you understand this alert message?	Understandability	All
In a real emergency, would this alert be relevant to you given your current situation and location?	Relevance	All
In a real emergency, would this alert annoy you?	Annoyance	All
In a real emergency, would this alert prompt you to take action?	Actionability	All
In a real emergency, would this alert prompt you to seek further information?	Milling behavior	All
Did this alert contain enough information?	Adequacy	All
Was the map sent with the alert useful?	Usefulness	Alerts with High-Information Map
Was the link embedded in this alert useful?	Usefulness	Alerts with External Link
In a real emergency, would you have preferred to receive this alert?	Hindsight relevance	Alerts with Geo-targeting, Location History and Location Prediction
This alert was targeted to participants who are in the alert region shown in the map, who have visited the region recently or who are moving towards the region. Would this alert be relevant to you in a real emergency?	Hindsight relevance	Feature group (Group B) of alerts with Location History and Location Prediction

At the end of each experiment, the WEA+ app solicited the subjects' overall impressions by a final, post-test questionnaire. Table 4.6 lists the questions included in the final questionnaire. Subjects responded to this questionnaire on their phones after the last alert of each experiment.

Table 4.6. Questions Used in Final Questionnaire in Experiments 1 and 2

Final Questionnaire Questions (Overall Perceptions)	
Question	Related Outcome Construct or Study Validity Indicator
Was the motivation of the trial clear to you?	Clarity of Motivation (Study Validity)
What did you think about the frequency of the alerts?	Frequency of Alerts (Study Validity)
How would you rate the realism of the alerts?	Realism (Study Validity)
How would you rate the intrusiveness of the trial?	Intrusiveness (Study Validity)

Final Questionnaire Questions (Overall Perceptions)	
Question	Related Outcome Construct or Study Validity Indicator
Do you believe wireless emergency alerts are useful?	WEA Benefits (Outcome Construct)
Do you believe wireless emergency alerts could save lives?	WEA Benefits (Outcome Construct)
Did inclusion of maps with some alerts increase their clarity and relevance?	Feature Preference (Outcome Construct)
Was the inclusion of links to external sources (URLs or social media tags) useful?	Feature Preference (Outcome Construct)
How would you rate the ease of use and installation of the WEA+ app?	Ease of Use (Study Validity)
How would you rate the effectiveness of the Trial Launch (Kick-off) event?	Kick-off Event Effectiveness (Study Validity—Experiment 1 only)
Some alerts targeted to a specific geographic area were delivered to you if you had recently visited the targeted area even if you were outside it when the alert was sent. Was this feature useful?	Feature Preference (Outcome Construct)
Some alerts targeted to a specific geographic area were delivered to you if you were moving towards the targeted area even if you were outside it when the alert was sent. Was this feature useful?	Feature Preference (Outcome Construct)
Enter any comments or suggestions you might have below.	Any

4.7. Design of Experiment 3

Experiment 3 was conducted with the iOS version of the WEA+ mobile app and focused on testing a single enhancement feature, but using a more complicated design that groups alert streams into incidents and a set of incidents into scenarios. It was also a single-factor randomized repeated-measures experiment, with a single binary factor, representing either the control (Group A) or the tested feature (Group B). Multiple alerts were sent over multiple days, as in Experiments 1 and 2. Figure 4.7 illustrates the design of Experiment 3.

The goal of Experiment 3 was to test a scenario-based digest view that supports situational awareness by overlaying information from streams of related messages. Each such stream represented a distinct, evolving emergency incident. The digest view on the iOS WEA+ app showed in a compact way only the latest digested information about each ongoing incident to the recipient, the last-updated incident being the default screen. This feature was compared to a control view that emulates ordinary WEA delivery as a single, interleaved sequence of messages from multiple incidents without any grouping, overlaying or consolidation.

The experiment was temporally sectioned into several scenarios as shown in Figure 4.7. A scenario included multiple parallel incidents. Each incident was a stream of messages, starting with an initial alert and continuing with various updates to emulate the evolving nature of the underlying emergency. The incidents within the same scenario could be independent or interrelated.

At the start of each scenario, Experiment 3 subjects were randomly assigned to either the control view (Group A) in which alerts were presented in a single stream similar to the current WEA service, or to the digest view (Group B), in which each ongoing incident had a summary screen compactly displaying the latest information about that incident. The group assignments were balanced, with subjects distributed evenly between the two groups.

At pre-determined points within a scenario, the iOS WEA+ app polled the subjects through a set of multiple-choice questions (Table 4.7) to assess their understanding of the overall situation or their situational awareness at that point in time. This assessment was transversal to all active incidents in the scenario. It was performed by asking the subjects about aspects of the state of the emergency situation that was the locus of their attention at the time they responded to the poll. The evaluated aspects were emergency type (nature), immediacy of the emergency, and what the subjects should do next (next action to be taken to respond to the emergency). In Figure 4.7, observations (answers) from a single poll are scattered along the time axis, because responses from subjects could arrive at any time within a scenario after a poll has been deployed. Since the scenarios could be evolving with multiple interleaved updates, answers to poll questions depended on the time of the response rather than the time of the poll.

Since each scenario had multiple parallel evolving incidents with multiple alerts spread over time, it was possible for an incident to receive a new update after a poll had been deployed, but before a subject had had the chance to respond to the poll. This potential misalignment between the timing of a mid-scenario poll and individual response delays meant that the correct answers depended on the time at which a subject responded to the poll rather than the time at which the poll was deployed. In addition, multiple parallel incidents implied that the subject could focus on one incident (possible the one with the most recent update) and answer the poll questions with respect to that incident. Such potential variance in the locus of attention meant that the correct answers also depended on the particular incident a subject happened to focus on when responding to a poll. Therefore, multiple correct answers were associated with each poll at each response point.

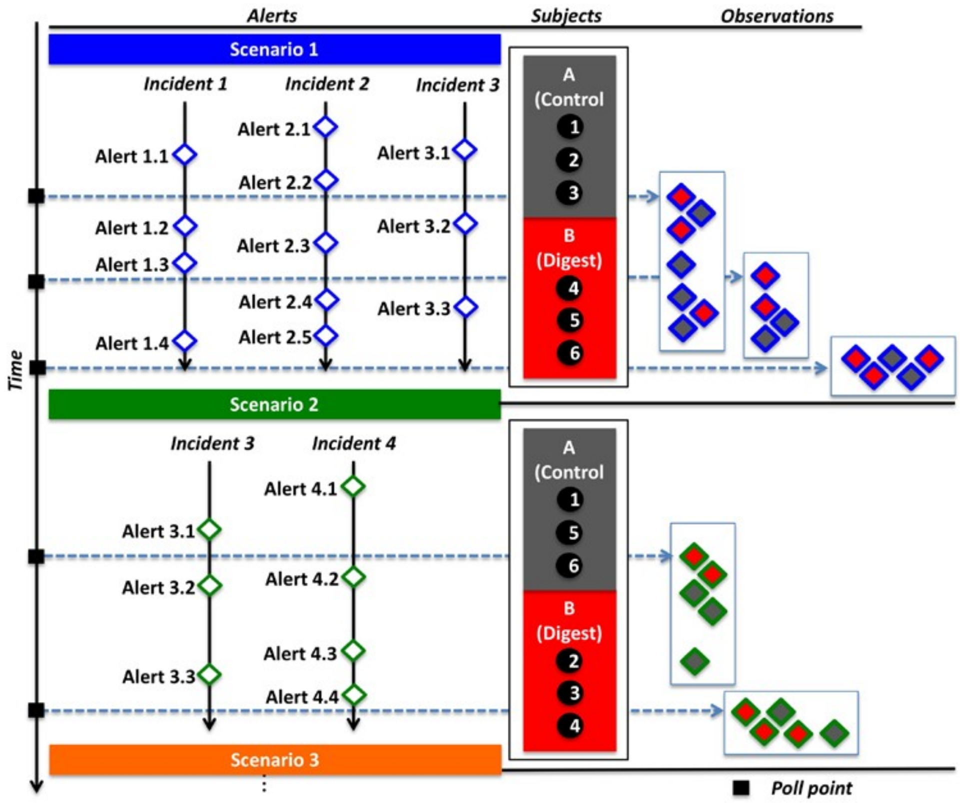


Figure 4.7. Experimental Design Used with iOS Subjects in Experiment 3 to Test Situation Digest Feature

If a subject responded to a mid-scenario poll, an observation was created for the subject-poll pair for the corresponding scenario. Certain questions were asked multiple times at different points in time within a single scenario to determine the users' situational awareness evolution with the progression of the scenario. Within each scenario, observations from polls belonging to that scenario were pooled for the same poll questions and Group A (control) and Group B (Situation Digest) observations for each poll question were compared. Observations were not pooled across different scenarios for an aggregate analysis because scenarios varied in complexity, rate of change, nature of change and theme to warrant different response trends. Totals over all scenarios for each poll question were reported for comparison purposes, however.

As in Experiments 1 and 2, Experiment 3 was concluded with a final, post-test questionnaire filled out by the subjects using the WEA+ app. The subjects were asked to answer a series of questions about their overall impressions regarding WEA in general and the experiment in particular, as well as their preferences between the regular WEA view and the incident digest view.

Table 4.8 lists the questions used in the final questionnaire and the corresponding outcome constructs or study validity indicators.

Experiment 3 was divided into five scenarios. Each scenario was designed to test how different factors, such as complexity, number of parallel incidents and rate of change, would impact the subjects' awareness of the situation at hand.

Table 4.7. Scenario-Based Poll Questions for Experiments 3 to Evaluate Situational Awareness

Poll Questionnaire - Mid Scenario	
Question	Related Outcome Construct
Given the information provided, what should you do now?	Situational Awareness (Type of Emergency)
An emergency is underway. What type of emergency is it?	Situational Awareness (Action)
An emergency is underway. When do you need to take action?	Situational Awareness (Immediacy)

Table 4.8. Questions Included in Final Questionnaire in Experiment 3

Final Questionnaire Questions (Overall Perceptions)	
Question	Related Outcome Construct or Study Validity Indicator
Was the motivation of the trial clear to you?	Clarity of Motivation (Study Validity)
What did you think about the frequency of the alerts?	Frequency of Alerts (Study Validity)
How would you rate the realism of the alerts?	Realism (Study Validity)
How would you rate the intrusiveness of the trial?	Intrusiveness (Study Validity)
Do you believe wireless emergency alerts are useful?	WEA Benefits (Outcome Construct)
Do you believe wireless emergency alerts could save lives?	WEA Benefits (Outcome Construct)
How would you rate the ease of use and installation of the WEA+ app?	Ease of Use (Study Validity)
You were presented with more than one view to represent the alert messages. One was a stream of messages and the other was an updating situation panel. Which one do you prefer?	Feature Preference (Outcome Construct)
Enter any comments or suggestions you might have below.	Any

4.8. Data Analysis Techniques

The majority of the trial data collected were responses to multiple-choice feedback questions and polls issued after alerts, in-between alerts or at the end of an experiment. These questions and polls were intended to measure a set of outcome constructs or dependent variables. The outcome construct levels represented in the questions and polls were of nominal scale (categorical), making them most amenable to an analysis based on frequencies. We therefore used the standard *Chi-square independence test* for all tested factors (with the null hypothesis that the tested factor is independent of the response distribution) and the *Chi-square goodness of fit test* to test responses about overall impressions that do not depend on the presence or absence of an alert feature or characteristic (Siegel 1994).

Appendix B2 provides further details of the data analysis approach.

In all Chi-square tests, the selected confidence level was 95 percent, corresponding to an alpha level of 1.5 for statistical significance. Thus we considered a test result to be *significant* when the p-value was below 0.05, rejecting the underlying null hypothesis.

We measured effect size in two different ways: (1) theoretical, using *Cramer's V* statistic for independence tests and the *Phi* statistic for goodness of fit tests; and (2) practical, *using odds ratio* for both kinds of tests (Zaiontz 2015). Given the underlying degrees of freedom, the theoretical effect sizes were interpreted as follows (Zaiontz 2015; Cohen 1992):

- Very small: Phi or Cramer's V smaller than 0.1
- Small: Phi or Cramer's V larger than or equal to 0.1 and less than 0.3
- Medium: Phi or Cramer's V larger than or equal to 0.3 and less than 0.5
- Large: Phi or Cramer's V larger than or equal to 0.5

An effect size between 0.2 and 0.3 as measured by Cramer's V or Phi is considered normal in studies dealing with human behavior where outcomes might be affected by multiple uncontrolled factors (Cohen 1992).

The odds of a positive response for a given group is defined as the ratio of the frequency of a positive response (example e.g., corresponding to the level "Yes" in some feedback responses) in the given group to the frequency of a negative response (e.g., corresponding to the levels "No" and "Unsure" together in some feedback responses) in that group. In all tested factors except Situation Digest, the mid-level responses are combined with the most negative level in calculating the odds ratio. In evaluating situational awareness, the middle response level (Partially Incorrect) is ignored in the calculation of the odds ratio and only the absolute positive level (Correct) and absolute negative level (Wrong) are considered. Based on the scheme adopted, the odds ratio for a tested factor (an alert feature or characteristic) is calculated as the odds of a positive response when the tested factor is present (the feature group or characteristic group) to the odds of a positive response when the tested factor is absent (the control group). Thus, the closer the odds ratio to 1, the smaller the effect size is. The interpretation of the odds ratio depends on the tested factor and its context.

The following subsections organize the findings by tested factor or outcome construct. For each tested factor or outcome construct, we first present the results for each applicable experiment separately. Then we present pooled results from multiple experiments when doing so is meaningful and sound. For Experiments 2 and 3 of Trial 2, observations from the two locations, Silicon Valley and Pittsburgh, are aggregated for analysis since the issued alerts were the same in content, type and timing, barring adaptations for local geography and local time. Location-specific results are reported and discussed only when location was found to be a significant factor. Data tables with detailed statistical analysis results are omitted here for brevity, but included in Appendices B3 to B6.

4.9. Results: Alert-Based Enhancement Features

Appendix B3 provides the detailed tabulated results for alert-based enhancement features tested in Experiments 1 and 2. A synopsis and interpretation are provided below for each tested feature.

Long Message

Synopsis. For Experiment 1, there were no differences in any of the outcomes between alert messages longer than 90 characters and alert messages obeying the current 90-character length limit. Both in general (between all long alerts and all short alerts) and pair-wise (comparing long and short versions of the same alert), longer messages naturally contained more information and detail about the underlying emergency than short messages.

The results were different in Experiment 2, in which messages were all longer than 90 characters. Instead we separated post-hoc the alerts into two length groups using the experiment's median alert length of 130 characters: alerts with more than 130 characters were treated as long and alerts with 130 characters or fewer were treated as short. There were no significant differences between the understanding and adequacy of short and long alerts according to this classification, but long alerts were found to be significantly more relevant and actionable and less annoying, with a small effect size in each case as measured by the Cramer's V statistic (between 0.14 and 0.16). Long alerts were also significantly more likely to cause milling behavior than short alerts, again with a small effect size. This finding may sound surprising because as one provides more info, the normal expectation would be reduced need to check elsewhere. However, we interpret milling behavior as evidence of interest in the information provided or evidence that people want to act or a pre-cursor to subsequent action (Hamilton et al. 2014). With respect to effect size measured by odds ratio, long alerts had at least 73 percent better odds of being relevant, actionable and prompting milling behavior than not and 49 percent worse odds of being perceived as annoying than not compared to short messages in Experiment 2. For Experiment 2, we reject the null hypothesis for the outcome constructs relevance, actionability, milling behavior and annoyance, and accept the alternative hypothesis that messages longer than 130 characters improve the associated outcome in the sampled population.

When observations from both experiments are combined in a pooled analysis, annoyance, milling behavior and actionability remain significant at an alpha level of 5 percent, but with reduced effect sizes, in favor of long messages (Cramer's V between 0.11 and 0.12). Relevance becomes significant only at an alpha level of 10 percent and with a very small effect size. Pooling of observations was sound because all included long alerts in Experiment 1, except one, were longer than 130 characters. We removed this data point from the pooled analysis to make the short-long classification consistent across the two experiments. Notably, when this single data point is also removed in Experiment 1 analysis, actionability difference becomes significant in favor of long messages as well.

Interpretation. Although message length does not appear to affect the recipients' understanding of an alert and perception of the adequacy of the information contained in an alert, when alert messages are long enough to allow more pertinent information to be communicated, the messages' relevance to the recipients, their actionability, and their chances of prompting milling behavior may improve. The significance and size of the improvement appear to be alert-specific. Thus alert specificity may be the reason for the differences in the results of the two experiments. Also, shorter messages have a larger tendency to annoy recipients than longer messages, possibly due to their reduced actionability in the eyes of the recipients. Pooled analysis from the two experiments suggests that actionability, milling behavior and annoyance are most affected by message length, in favor of long messages. Alert messages that contained information worth over 130 characters yielded significant differences in the pooled analysis for these outcomes. It is a bit surprising that while actionability and annoyance differed between two message-length groups, there were no perceived differences in information content (adequacy). This discrepancy is curious, and might be due to the ambiguity of, and variances in, the interpretation of the term "enough" in the adequacy question "Does this alert contain enough information?"

High-Information Map

Synopsis. Certain alerts were displayed to a random sub-sample of the recipients with a map of the alert area showing the geo-target polygon and the location of the recipient. Hamilton et al. (2014) refer to such maps as "high-information." We compared alerts displaying a high-information map to those that did not. Alerts containing extra enhancement features that could confound the comparison were excluded from the analysis.

In Experiment 1, alerts with a high-information map were found to be significantly more relevant by the recipients compared to alerts without maps. The effect size was small as measured by Cramer's V.

Although Cramer's V was low, the odds ratio indicated that alerts with high-information maps improved the odds of finding an alert relevant by 75 percent.

The remaining outcome constructs did not exhibit any other significant differences in Experiment 1; however, milling behavior and adequacy were not reliably measured due to a low number of observations (all observations for these two constructs were associated with a single alert and do not satisfy the minimum sample size requirement of the Chi-square independence test). Therefore, we cannot draw any conclusions about the effect of maps on milling behavior and perception of adequacy for Experiment 1.

Of those subjects who received high-information maps, 79 percent indicated that they found the map useful. This result is highly significant relative to the null hypothesis of a uniform distribution (we assumed that if subjects were indifferent to the usefulness of maps, the answers measuring this construct would be distributed uniformly).

Relevance and adequacy was significant only at an alpha level of 10 percent in Experiment 2, but usefulness remained highly significant with nearly 85 percent of the subjects who received maps finding them useful.

Combining the data from Experiment 1 and 2 in a pooled analysis, we observe that high-information maps significantly improve alert relevance, adequacy and usefulness at the chosen alpha level of 0.05. The magnitude of the improvement remains small according to Cramer's V statistic for both constructs. The odds ratio is still considerable for relevance at 1.35, indicating a 35 percent improvement, but lower than in Experiment 1 (1.75). Combined usefulness is highly significant with a large effect size.

In the final questionnaire of both experiments, the subjects were asked, based on their experience with alerts delivered with and without maps during the respective trial period, about their impressions of the map feature. Ninety percent of Experiment 1 subjects and 87 percent of Experiment 2 subjects responded that maps improved the clarity and relevance of alert messages. These results were also highly significant.

Interpretation. Overall alerts displaying high-information maps increase the relevance of the alerts to the recipients. The improvement is significant. The magnitude of the improvement is theoretically small according to Cramer's V, but according to odds ratio, a more practical effect size measure, maps improve the odds of finding an alert relevant considerably (by 75 percent). When asked explicitly whether they found the map displayed with the alert just received useful, subjects overwhelmingly responded positively. This result is further supported and strengthened by the final questionnaire responses. We conclude that high-information maps have a positive effect on the relevance of an alert conditional on the recipient's location and situation, and are a highly desired WEA enhancement feature. Maps also appear to affect the information content of alert messages, as measured by adequacy. The improvement is both smaller and subtler than for relevance, however.

We did not find any evidence of a positive or negative effect of high-information maps specifically on actionability, annoyance and milling behavior.

High-information maps work synergistically with geo-targeting and other geo-targeting-based enhancements. These enhancements are easily implementable on the phone if alerts carry their geo-target. Furthermore, next generation smart phones are expected to feature pre-cached maps. Therefore, inclusion of high-information maps is well worth considering in future WEA implementation.

Geo-targeting

Synopsis. Both trials were conducted in the relatively small and concentrated metropolitan areas of Pittsburgh (covering primarily central Pittsburgh and surrounding areas) and Silicon Valley (covering primarily Mountain View and surrounding cities of Santa Clara county), in which our subject populations were concentrated. The geographical distribution of the subject populations was suitable for assessing the effect of high-resolution, precise geo-targeting with highly-localized alerts. Therefore, our results do not generalize to coarser resolutions of targeting (e.g., state-level and larger).

In the study's context, geo-targeting has a specific meaning: it refers to the precise, fine-grained type of targeting meant to augment the standard targeting mechanism used in the standard WEA service.

We issued several geographically-targeted alerts to areas within the coverage areas of the trial, defined by polygons of varying sizes (from a few city blocks to spanning multiple neighborhoods) and shapes (elongated, square-like, convex, concave, with varying number of vertices up to 15). We compared the reaction to these alerts from the subjects who happened to be in the alert area or entered the alert area before an alert's expiration to those sent to the whole subject population (no filtering based on an embedded geo-target).

In both experiments, the subjects found geo-targeted alerts significantly more relevant than non geo-targeted alerts. The effect size ranged from medium-small to medium (with Cramer's V equal to 0.32 in Experiment 1, 0.22 in Experiment 2 and 0.27 overall). Odds ratios were substantial: 6.11 for Experiment 1, 2.25 for Experiment 2 and 4.12 overall (Experiment 1 and 2 combined).

Adequacy was significant only for Experiment 1 in favor of geo-targeted alerts with a small effect size. It was not significant in Experiment 2, and across the two experiments, significant at an alpha level of 10 percent. Overall, odds ratio was still reasonably high at 1.62.

For all geo-targeted alerts, shortly after an alert's expiration, we showed the alert to the subjects who were outside the alert area during the alert's period; having not received the alert, we asked them whether they would have preferred to receive that alert. In Experiment 1, 70 percent said they would have. This number was lower at 57 percent in Experiment 2 and 63 percent combined. We refer to this after-the-fact response as hindsight relevance. The lower hindsight relevance is, the more effective geo-targeting is.

The remaining outcome constructs, understanding, actionability and milling behavior, were also better with geo-targeting with odds ratios ranging from 1.21 for milling behavior to 2.30 for understandability, but not significant except actionability in the pooled analysis, which was significant at an alpha level of 5 percent. The lack of significance in the separate analysis is due to insufficient sample size. The odds ratio for actionability after observations from both experiments have been combined was 1.75, representing a 75 percent improvement in the odds of the recipients' finding an alert relevant with geo-targeting.

Interpretation. Precise, fine-grained geo-targeting improved alert relevance to recipients as hypothesized. The improvement was highly significant with a near-medium theoretical effect size as measured by Cramer's V and considerably large practical effect size as measured by odds ratio (over three times improvement in the odds). Hindsight relevance, an alert's relevance to those outside the alert's geo-target, was not as low as expected at 63 percent; this result is attributed to the relatively small size of the experiments' geographical coverage areas. Actionability was also better with geo-targeting, but this effect was not as strong, and it was prominent only when the data from both experiments were pooled. We conclude that geo-targeting has a significantly positive impact on alert relevance and a small to moderate impact on actionability.

We expected hindsight relevance to be lower than 50 percent, however, the experiments' total coverage areas were probably not large enough to cancel the effect of location history, people having an interest in multiple districts routinely visited (for example where their workplaces, residences, shopping areas and children's schools are located) regardless of their current location. The analysis of the Location History feature supports this explanation.

The effect of geo-targeting on alert relevance constitutes the strongest finding of the PUT phase.

Location History

Synopsis. The Location History feature was evaluated as a part of Experiment 2, in the relatively small and concentrated metropolitan areas of Pittsburgh (covering primarily central Pittsburgh and surrounding areas closer to CMU campus) and Silicon Valley (covering primarily Mountain View and surrounding cities of Santa Clara county). This feature was used in conjunction with geo-targeting. As with the Geo-targeting feature, the particular geographical distribution of the subject populations implies that we cannot generalize the results to coarser targeting resolutions than used in the Experiment 2.

Throughout the duration of trial, the WEA+ app kept recording the geo-locations of mobile phones every five minutes. This record constituted the location history of a user. Since the trial lasted only a week, we did not use a sliding history window. When a geo-targeted alert was sent with the Location History feature turned on, the WEA+ app used the local location history on the phone to determine if the recipient had ever visited the alert region using the geo-target polygon included with the alert. If the recipient had visited the alert area in the past or was already inside the alert area, the alert was shown; otherwise it was discarded. As with the Geo-targeting feature, we compared the reaction to these alerts from the subjects who received these alerts to the reaction to alerts who were sent to the whole subject population with no filtering.

Subjects found geo-targeted alerts filtered with Location History to be significantly more relevant than alerts that were not geo-filtered using any scheme. For this outcome construct, a Cramer's V of 0.25 indicates a small to medium effect. An odds ratio of 335 percent in favor of Location History suggests that this feature was highly desirable. Alerts filtered with Location History were also significantly more actionable. For actionability, a Cramer's V of 0.19 indicates a small effect size. Finally, alerts filtered using this feature were significantly more adequate, with a Cramer's V of 0.15, indicating a small effect. No significant differences were observed for understanding, annoyance and milling behavior, however.

Hindsight relevance was also measured. Subjects who did not receive these alerts as a result of location-history based filtering were asked if they would have preferred to receive them. Sixty-eight percent said they would have preferred to receive the alert. Recall that the lower hindsight relevance is, the more effective the underlying filtering scheme is. As with pure geo-targeting, we expected hindsight relevance to be lower than 50 percent. Again, the experiment's total coverage areas were probably not large enough to cancel the possible bias of location proximity.

The usefulness of the Location History feature was assessed in the final questionnaires deployed to all subjects at the end of Experiment 1. Overall, 79 percent of the subjects found the Location History feature useful. This finding was highly significant.

Interpretation. We expect filtering an alert based on the recipient's location history to improve an alert's relevance to recipients. In the trial, the improvement was highly significant with a near-medium theoretical effect size as measured by Cramer's V (0.25) and considerably large practical effect size as measured by odds ratio (over four-fold improvement in the odds). Actionability and adequacy are also expected to improve when geo-targeting is combined with location-history-based filtering. Hindsight relevance, an alert's relevance to those outside the alert's geo-target, was not as low as expected; however, this result could be attributed to the relatively small size of the experiments geographical coverage areas. We conclude that Location History is a highly desirable feature and is likely to strengthen the positive impact of geo-targeting, in particular on alert relevance and actionability. The feature can be implemented on the user's device while respecting the user's privacy since location data is not sent outside and all filtering is performed locally.

Some of the parameters used in the Location History feature were the length of the history window (unlimited in the trial) and threshold visit frequency (set to 1 in the trial). We do not know how these parameters affect this feature's effectiveness. Longer-duration studies are needed to investigate the significance and optimal choice of these parameters.

Location Prediction

Synopsis. The Location Prediction feature was also evaluated as a part of Experiment 2 as an additional adjunct to geo-target-based filtering, subject to the same contextual factors and constraints as the Location History feature regarding the geographical distribution of the subject population and applicable targeting resolution. We do not repeat them here.

In this feature, when the WEA+ app receives a geo-targeted alert with the Location Prediction feature turned on, the app first uses its local location history database (same record used in the Location History feature) to retrieve the user's last three known locations. Based on the geo-coordinates of these locations, the app then predicts the recipient's possible future locations. Since an exact trajectory is difficult to calculate, considering bends and changes in the unknown route of a user in motion, the app assumes a straight-line trajectory with some margin around it. This results in an expected area of future locations, which in conjunction with the user's estimated speed and the specification of the geo-target, is subsequently used to determine if the recipient could enter the alert area within the alert's active period. If the prediction is positive, the alert is shown; otherwise the alert is discarded.

Note that here we measured the effectiveness of a specific location prediction algorithm. We consider this algorithm as representative of a simple, but plausible scheme. The results are clearly predicated on the scheme used for location prediction, and with this point in mind, they should be considered at face value.

As done in Location History, we compared the recipients' reactions to these alerts to reactions to alerts that were sent to the whole subject population with no filtering. None of the five studied outcome constructs (understanding, relevance, annoyance, actionability, milling behavior and adequacy) was significantly different for the Location Prediction feature compared to the control group. Effect sizes, as measured by Cramer's V, for all constructs were small in favor of the feature. Absence of significant results may be due to the small sample size used for testing this feature or due to the imperfection in prediction algorithm.

Hindsight relevance was also evaluated by follow-up alerts sent to the subjects who had not previously received an alert with the Location Prediction feature. The follow up alert asked these subjects if they would have preferred to receive the original alerts. Sixty-six percent said they would have preferred to receive the alerts even though they were outside the alert region and not determined to be moving toward it. This response was significant. As in Location History, we had expected hindsight relevance to be lower, preferably below 50 percent; the relatively small size of the geographical area covered may have made most alerts relevant to most subjects, increasing hindsight relevance. The location prediction algorithm might also have missed some true positives by failing to sufficiently accurately determine the future location of a subject within the active period of an alert.

The usefulness of the Location Prediction feature (as implemented by a simple prediction scheme) was assessed in the final questionnaire deployed to all subjects at the end of Experiment 2. Overall, 76.6 percent of the subjects found geo-targeting augmented with location prediction useful. This finding was highly significant.

Interpretation. We are not able to draw any strong conclusions regarding the augmented filtering of geo-targeted alerts through location prediction on the recipient phone. Beyond the subjects' perception of the usefulness of this feature, there was no significant evidence that it improved any of the specific outcome constructs evaluated, possibly due to the small sample size. Also location prediction, by definition, is approximate; it is subject to great uncertainty. Absence of significant results may have been due to the small sample size used for testing this feature or due to the imperfection in prediction algorithm.

The algorithm used in the feature was rough and relied only on simple contextual information. Our results are valid for this scheme and similar schemes with comparable accuracy. The results may get better and stronger with more advanced schemes. For example, it is possible to improve prediction accuracy by more sophisticated use of location history, machine learning techniques, and taking advantage of activity recognition (e.g., whether a user is driving, cycling, walking or stationary can be factored in, and a location change is predicted only if a user is in motion at a sufficiently high speed relative to the proximity of the alert target to the user's current location.)

In conclusion, the use of location prediction in WEA delivery could be valuable, but deserves further study before a recommendation can be made.

External Link

Synopsis. In Experiments 1 and 2, selected alerts were issued with clickable website links or Twitter hash tags to test whether inclusion of such external references containing additional information about the underlying emergency influenced any of the outcome constructs. We did not find any notable differences, neither in terms of magnitude of the effect nor in terms of statistical significance, when we analyzed the two experiments separately. Some questions had too few answers in either experiment to reveal significant results. Therefore, we do not report the detailed results of the experiments separately for this tested feature. For example, in Experiment 1, the questions measuring adequacy and milling behavior had a total of only 18 observations to cover the three outcome levels for this construct and in Experiment 2, certain outcome construct levels produced fewer than six observations.

Combining the results from the two experiments fortunately yields enough observations. In the joint analysis pooling the data from both experiments, adequacy was significant in favor of alerts containing external links with a small effect size as measured by Cramer's V (0.11) and an odds ratio of 2.10 (110 percent improvement in the odds of better adequacy for alerts with external links). The other outcome constructs did not produce any significant differences.

When the recipients were asked whether they found the external link in the alert that they had just received useful, a 54 percent majority, responded positively. This figure is the combined statistic for Experiments 1 and 2; it was 49 percent for Experiment 1 and 69 percent for Experiment 2 (Experiment 1 had four levels for this question, including an additional answer "Did not notice," which was combined with the answer "Unsure" in the pooled analysis). The results were highly significant for both experiments, as well as the pooled analysis.

The usefulness result was re-confirmed in the final questionnaires deployed to all subjects at the end of the experiments. Overall 67 percent of the subjects from both experiments (75 percent in Experiment 1 and 64 percent in Experiment 2) indicated they found the external links useful when they were included in the alerts. These findings were highly significant.

Interpretation. Inclusion of external links have been proposed as a way to compensate for WEA messages' supposed inability to convey sufficient information due to their length limitation. Clickable links to external Internet resources, however, are not currently permitted due to concerns for network congestion and possible subsequent service interruption potentially caused by simultaneous access to those resources by a large number of smart phones in a small area. Based on the pooled results from both experiments, we conclude that inclusion of external links in the form of URLs to relevant websites or social media tags are a highly desired feature even if their impact on understandability, relevance, milling behavior and actionability is unclear. External links may improve the information content of alert messages by allowing the subjects to easily access additional resources directly from their phones.

If potential threats to network stability can be adequately addressed and the message length limit accommodates the addition of shortened URLs and hash tags into the alert text, future WEA service implementations should explore permitting clickable external links in alert messages. As was suggested by one respondent in the AORS survey comments, a link could be to a short web page locally cached at the cell tower, which would alleviate network stability concerns. LTE capabilities similarly could render such concerns moot. We are unable to make a more definite recommendation based on the evidence collected.

Text-to-Speech (TTS)

Synopsis. We tested TTS as a feature to supplement alert notification with a spoken version of the alert text. Normally, recipients are notified of the delivery of an alert through vibration and a special ringtone, if these capabilities are enabled on their phones. At the beginning of both trials, the subjects were asked to enable audio and vibration on their phones and keep the media volume at an easily audible level, but we had no way of ensuring that these instructions were followed by all participants.

In Experiment 1, understanding, actionability and annoyance were significantly improved with TTS alerts, but the magnitude of the effects were small, ranging from 0.16 for actionability to 0.20 for understanding as measured by Cramer's V. Odds ratios were higher, ranging from 1.68 to 3.08 for these outcome constructs. However, the same effects were not observed in Experiment 2; no significant differences were noted in any of the outcome constructs. In the pooled analysis, only understandability and actionability were significant at an alpha level of 5 percent, and Cramer's V value was miniscule to small at 0.09 and 0.10, respectively for understandability and actionability. Odds ratios also dropped significantly to 1.79 and 1.22, respectively. Surprisingly, in Experiment 1, nearly half as many subjects in the TTS group found alerts annoying as compared to those in the non-TTS group.

Did spoken alerts delivered with TTS influence the subjects' rate of response and speed of reaction? For example, did such alerts cause the subjects to delay viewing the alerts on their phones because they were better able to assess alert relevance based on the spoken alert text alone? In Experiment 2, we measured each subject's response delay to each alert to answer this question. Response delay was defined as the elapsed time in seconds between the interception of an incoming alert by the WEA+ app on the recipient's phone and the opening (and subsequent viewing) of the alert by the recipient by hitting a button on the notification screen. We measured average response delay for TTS and non-TTS alerts with different caps on the response delay (see Appendix B3). The data were sectioned depending on an upper response delay limit. Except when no cap was used, response delays were consistently longer for TTS alerts. When there was no limit (all observations), response delay was longer for non-TTS alerts. When no cap was used, the average delay was significantly skewed due to observations with very long response delays. We did not test the statistical significance of the differences in average response delay because the observations were extremely irregular with no readily identifiable patterns.

Remarkably, response rate, both in terms of eventually viewing an alert and in terms of eventually giving feedback about an alert, was significantly higher for TTS alerts with a medium effect size. Thus TTS alerts were more effective in prompting an eventual reaction.

Interpretation. We did not find compelling evidence of any differences in understandability, adequacy, actionability, milling behavior and relevance with the use of TTS to amplify alert notification. TTS might have a small effect on understandability, but uncontrolled confounding factors and design differences between the experiments prevent us from drawing any strong conclusions. Annoyance surprisingly was worse in Experiment 1, however annoyance is likely to be affected by demographics as well as temporal and other contextual factors.

TTS-enhanced alerts appear to delay average short-term response and accelerate average long-term response, the former possibly due to the recipients' ability to assess an incoming alert via spoken text without having to open the alert on the receiving device. The reason behind the impact on the long-term response delay is unclear at this point, but may be attributed to improved ability to remember the arrival of an alert long after the notification time.

TTS alerts also improved eventual response rate, which may be interpreted as a precursor to *real* actionability, complementing *hypothetical* actionability (the actionability outcome construct) measured by the feedback question, “Would this alert prompt you to take action?” Again, better retention and the convenience of assessing the alert without having to check the phone might be the reasons for the higher response rate of TTS alerts. Combining this result with the result of Experiment 1 suggests TTS could be beneficial for actionability in proper contexts (e.g., driving).

We attribute the differences between the results of the two experiments to the specifics of the alerts used, confounding of the use of TTS by other factors such as alert timing, inability to control the media volume and differences in sampled populations. It is possible that more subjects turned off or reduced the media volume of their phones, which would effectively disable TTS. Moreover, in Experiment 2, we did not specifically match TTS alerts with equivalent non-TTS control alerts sent at the same time using randomized grouping. Instead we compared responses to the whole set of alerts sent with only TTS enabled and all other features disabled, with the set of all alerts having no features enabled (subsets of this latter set doubled as control for other tested features). This design difference and the resulting interference from uncontrolled factors limit the internal validity of the results of Experiment 2 and the pooled analysis.

We expected TTS to increase the annoyance rate because spoken alerts could be perceived as more intrusive. We observed the reverse, however. This reversal may be attributed to the interactions between the presence of TTS, the sample population, alert timing and experimental context. The majority of the participants to Experiment 1 were graduate students with similar schedules (since they were part of a small campus), and daytime alerts avoided TTS to prevent interference with classes, but non-TTS alerts could be received during class hours. Possibly, alerts received during class hours, which tended to be non-TTS alerts, were perceived to be more intrusive and annoying. Alert timing analysis for Experiment 1 is consistent with this explanation; annoyance rate was higher for daytime alerts than for off-hour alerts (although this effect was not significant), which was not the case in Experiment 2. Experiment 2's sample had different demographics; student representation was nearly half of that of Experiment 1 and most of the student subjects, being part of a large campus in Pittsburgh, likely had more heterogeneous schedules on average than those who participated in Experiment 1.

With regard to average response delay, we interpreted a delay exceeding five minutes as motivated by convenience over urgency. This behavior may have played a role in the longer average response delay (considering observations from both experiments) for non-TTS alerts, which had a significantly lower response rate. Non-TTS recipients did not have as strong a tendency as TTS recipients to view an alert and respond to it, and when they did, they appear to have taken their time.

Since TTS is relatively easy to implement and may usefully increase response rate, it is a feature deserving further consideration and study.

4.10. Results: Alert Characteristics

In Experiments 1 and 2, we tested two factors pertaining to the attributes or nature of the alerts themselves rather than being possible enhancements or changes to the WEA service. We discuss the findings related to these factors below. Tabulated results and details of statistical analyses are given in Appendix B4.

Alert Timing

Synopsis. Alert timing refers to the period during which an alert is active, defined by an issue date-time and an expiration date-time. A recipient is notified when an alert becomes active and as soon as the recipient's phone determines if the recipient is in the alert's target area or has entered the alert area before the alert expires. To be on the safe side, the recipient is notified if the device cannot determine the recipient's location (for example when the receiving phone's location services are disabled).

In Experiments 1 and 2, we specifically tested whether there were any differences in the subjects' reactions to alerts sent during the day (between 7 a.m. and 7 p.m.) when subjects are more likely to be preoccupied with work/school responsibilities and daily chores (tagged as day alerts) and alerts sent during the evening hours and later at night (after 7 p.m. to midnight) when subjects are more likely to be less preoccupied (tagged as night alerts). We did not issue any alerts after midnight. Because of the interaction between the TTS feature and alert timing (TTS was used more frequently in night alerts than in day alerts), we filtered non-TTS alerts out, thus settling for a smaller sample size for the alert timing analysis. Alternatively, we could have excluded TTS alerts, but there were not sufficiently many non-TTS night alerts for a viable analysis. A smaller sample size in turn is less likely to reveal differences between the compared groups; however, our initial unfiltered analyzes clearly demonstrated that failing to control for TTS could yield misleading alert timing findings.

The results from the two experiments and from the two locations in Experiment 1 were mixed and not completely consistent.

Experiment 1 subjects were predominantly graduate students studying in a small campus. They had largely overlapping schedules and likely lived close to the campus (we did not collect residential information from the participants to confirm this conjecture, however). We found no difference between day and night alerts for this sample, except for adequacy, which was significant in favor of night alerts, and the effect size, which was close to medium (Cramer's V = 0.28).

Experiment 2 subjects had split demographics in two different locations. Pittsburgh subjects were mostly students (less than 15 percent were non-student researchers) studying in a major campus. About half of Silicon-Valley subjects were students and staff at CMU's Silicon Valley campus and the rest were mostly government employees and CERT volunteers with a connection to emergency services. For this sample overall, there were no notable differences between day and night alerts, including for adequacy. When we sectioned the data according to location (Pittsburgh vs. Silicon Valley), the results were not different.

Differences in adequacy, as well as in other outcomes, were insignificant in the pooled analysis.

Interpretation. Based on the small sample sizes, interaction with other tested features, complexity of the demographic and alert-specific confounding factors, and inconsistency of the results regarding adequacy (only significant finding in Experiment 1), we are unable to draw any conclusions about the effect of alert timing on the evaluated outcome constructs. Alert timing would be best tested with a more specialized experimental design better suited for isolating it.

Alert timing is particularly difficult to isolate partly because during off-hours, the sampled population tends to be more dispersed compared to during work hours and dispersed differently depending on the sample. Additionally, demographic factors may have too much influence when uncontrolled. When the population is more diversely dispersed, extra alert-specific attributes (e.g., type and affected area) may skew the results. Demographic factors (students vs. working/stay-home/retired adults, small-campus students with similar schedules and who tend to live close to the campus vs. large-campus students with diverse schedules and who tend to be more spread at night in a metropolitan area with better public transport) and contextual factors may also change behavior disproportionately during day and night and consequently bias the results. Possibly, subjects in the Experiment 1 sample were able to pay more attention to alerts received after 7 p.m. during which they were less pre-occupied. This might have allowed them to better assess the information content of incoming night alerts.

Update Alert

Synopsis. Experiment 1 also tested whether alerts that updated the status of an emergency communicated with a previous alert caused a different reaction than the initial alerts. The sample size was very small; three pairs of update alerts were issued, with one pair radically changing the action of a previous alert, one pair maintaining the action and one pair cancelling it. There were no significant differences between understanding, relevance and annoyance. Actionability, milling behavior and adequacy were statistically different for update alerts as compared to initial alerts, however. Cramer's V was 0.19 for actionability and adequacy (considered a small effect) and 0.25 for milling behavior (considered a small-medium effect). Actionability and milling behavior were both worse with update alerts, with odds ratios of 0.47

and 0.50, respectively, meaning initial alerts had nearly twice the odds of being actionable and prompting milling behavior than update alerts.

Adequacy was better for update alerts than for initial alerts (Cramer's V = 0.19, odds ratio = 1.89). Therefore, subjects were more likely to think that update alerts contained enough information than they were to think that initial alerts contained enough information.

Interpretation. Initial alerts are also more likely to trigger information-seeking behavior as compared to update alerts. However, update alerts appear to increase the information content of an ongoing emergency; they were more likely to be perceived as containing sufficient information compared to initial alerts. This behavior of update alerts is positive, but it is conditional on whether they correctly affect the recipient's situational awareness. The standard WEA service treats update alerts in an ad-hoc manner, and this may lead to confusion in complex, rapidly changing situations. We view this potential threat as a motivator for new schemes that address situational awareness in a more structured way, as done with the Situation Digest feature.

Higher adequacy of update alerts is consistent with the opposite effect on milling behavior. If recipients think they have more information, they would be less likely to seek additional information.

We are not able to interpret the differences in actionability due to the small sample size and low diversity of the factor group. Actionability is too alert-specific and only one pair of the update alerts issued modified a previously specified action in any significant way, which is precisely when actionability matters most. Responses to the remaining two pairs of alerts likely skewed actionability unfavorably for update alerts.

4.11. Results: Situational Awareness

Situational awareness was studied in Experiment 3 using a more sophisticated approach than other outcome constructs studied in Experiments 1 and 2. It was evaluated by multiple polls deployed at pre-determined times between alerts during separate, evolving scenarios of varying complexity. Rather than being based on subjects' perceptions and self-assessment, situational awareness was measured by the correctness of the answers given to polls. Therefore it was measured in a more objective manner than the other outcome constructs. A poll's correct answers were time- and scenario-dependent. Correctness was directly tied to the subjects' confusion or understanding of a scenario. At any given time in a scenario, a subject's situational awareness was deemed high to the extent that the subject answered the corresponding poll questions correctly.

Since a scenario could contain multiple interleaved incidents evolving simultaneously, each poll question could have multiple correct answers at any given time. Depending on the choices selected in the answer, each question was graded as wrong, partially correct or correct. If the subject selected an answer including at least one incorrect selection, or a selection representing an assumption clashing with the last updated information about an incident, the question was graded as Wrong. A Wrong classification was equated with confusion or lack of understanding. If the subject selected some of the correct answers out of the set of all correct answers, but not *all***,** then the question was graded as Partially Correct, only if the subject's answers did not include an incorrect selection. Finally, if all of the possible correct answers were selected with no incorrect selections, the question was graded as Correct.

Through the poll questions, the subjects were evaluated on their assessment of three main aspects of the last known status of a scenario: the nature, or *type*, of an underlying emergency, the *action* to be performed as a consequence of an underlying emergency, and the *immediacy* of an underlying emergency. Recall that a scenario could involve multiple incidents, each with a different type, action and immediacy.

The three aspects — type, action and immediacy — constitute the main components of the situational awareness construct for our purposes. Situational awareness outcomes were compared for two different subject groups within each scenario: the control group who were presented with a normal WEA view (flat list of interleaved alert messages) and the feature group who were presented with the digest view (consolidated and incident-segregated view implementing the Situation Digest feature on the WEA+ app). Our hypothesis: the digest view would improve situational awareness for complex scenarios.

Next we discuss the findings regarding each of the three aspects for each scenario. Tabulated results are provided in Appendix B5. For completeness, Appendix B5 also reports pooled results from all scenarios for each aspect, the pooled results are not necessarily meaningful on their own because the scenarios differ significantly in complexity and nature. After giving a synopsis and interpretation of each aspect, we provide overall conclusions regarding situational awareness.

Type of Emergency

Synopsis. Measuring the users' understanding of the nature of the emergency when faced with multiple unfolding incidents was important to determine how easily a user could get confused when presented with potentially conflicting information, resulting in a wrong user perception of the situation.

We were able to show that in the complex scenarios — Earthquake with Plume, Random Alerts and Alien Catastrophe — the digest view performed significantly better compared to the regular WEA view. In terms of effect size, Cramer's V ranged from 0.23 to 0.31, considered small to medium in these scenarios. The odds ratios ranged from 2.98 to 27.23; the digest view improved the odds being correct approximately three to 27 times in terms of the subjects' understanding of the type of the emergency.

In the severe weather and bad weather scenarios, the type of the alerts issued barely changed. In particular, all the alerts issued in the latter scenario were of the same type. Thus the subjects were not likely to get confused regarding the type of emergency. Consequently, no differences were observed between the two groups in the Bad Weather scenario. The Severe Weather scenario had a single interleaved alert of a different type, a traffic alert. In this case, there was a slight difference between the groups, but not enough to be significant.

Interpretation. As the complexity of a scenario increases with respect to type of emergency, Situation Digest performs increasingly better. For complex scenarios, the improvement in the situational awareness outcome with respect to type of emergency is significant for this feature.

Action to Perform

Synopsis. Since the alerts issued during the experiment were not related to real emergencies, we had no way to follow up about the actual actions taken by the subjects in reaction to these alerts. To measure the subjects' level of understanding and awareness in terms of what actions they had to take upon receiving an alert, we instead asked them to choose the correct actions from a list of possible actions at multiple points within each scenario. Only a subset of the provided choices was aligned with the last known information about the scenario.

The first scenario, Earthquake with Plume, was designed such that the instructions given in the alerts did not change much. Thus, the cognitive load required by the subjects to re-assess the situation after a new alert was little. Even this small change in the instructions was sufficient to trigger different behaviors in the groups, however. The digest view had a positive effect on the users understanding of what to do. Furthermore, the odds of this group to be correct in their assessment of the situation were more than twice (odds ratio = 2.20) as those of the control group. Similar results were observed for the second scenario, Random Alerts. This scenario involved multiple instructions given for different alerts, but the alerts were completely independent, belonging to separate unrelated incidents. There was no change in the information over time on a per-incident basis and the incidents did not interact. Although neither of these two scenarios was significant, there were still significant differences between the two groups, in favor of the digest view.

The Severe Weather scenario was designed to test the increasing complexity of an incident with multiple changes to the actions to take. In this scenario, there is a significant difference between the groups, in favor of the digest view. The feature group was approximately five times (odds ratio = 4.73) more likely to get the action to take right than wrong compared to the control group.

The longest and more complex scenario of the experiment, Alien Catastrophe, had similar results as the previous scenario, favoring the digest view over the normal WEA view. The effect became more pronounced as the rate of change of the scenario increased. Although Cramer's V in this case was small (0.16), the odds ratio was considerable (5.55).

Finally, the last scenario, Bad Weather, was designed to simulate the simplest of cases where only a couple of alerts are issued in regards to a single incident. This was done to assess if the digest view would hurt the user's situational awareness in a simple non-evolving emergency. The results suggest that this view did not compromise situational awareness with respect to action to take. There was even a slight improvement in the feature group, but the difference was not significant.

Interpretation. As the complexity of an emergency scenario increases with respect to the prescribed action to perform, the Situation Digest feature performs increasingly better. For complex scenarios, the improvement in situational awareness with respect to this aspect is significant. This advantage of Situation Digest is probably due to the increased cognitive load of having to re-construct a changing situation in memory in the standard WEA view, as opposed to simply viewing it already re-constructed on the screen.

Immediacy of Emergency

Synopsis. The immediacy of an emergency was conveyed with each alert in the form of an implicit temporal clue ("X is expected," "Y is imminent," "Z is over") or expression of uncertainty ("Z is possible") in the normal WEA view. The digest view also had explicit visual and verbal cues in the form of an imperative ("Execute Now," "Be Prepared," etc.). In this analysis, we wanted to assess how the changing nature of the information influences the perception of immediacy for the user by asking the subjects when they needed to take action during the polls.

Only the first scenario resulted in a significant difference in the responses of the two groups (Cramer's V = 0.28, odds ratio = 3.19). The other scenarios exhibited no significant differences.

Therefore, the immediacy aspect of situational awareness did not quite behave like the other aspects (action and type). Further analysis revealed that when the subjects were asked when they needed to take action, a great majority answered by selecting "Immediately" from the available options whenever an explicit action was specified. The other options available were "In the near future," "Eventually, but not yet," "I don't need to take action," and "I am not sure," but they were rarely chosen. The subjects behaved in this way independent of the scenario or timing of the poll. It seems they assumed that every emergency and every new alert necessitates immediate action. This assumption is not entirely unreasonable; after all, a warning system with the kind of reach of the WEA service is supposed to mobilize the public affected by an emergency and every alert and update is automatically interpreted as urgent. This automatic uniform interpretation accounts for any differences in the results; in the cases where "Immediately" did not happen to be the right response and immediacy was more explicit in one view versus the other, the differences were significant; otherwise, they were not.

Interpretation. We did not find any strong and consistent evidence that Situation Digest affected the recipients' true understanding of the changing immediacy of an ongoing emergency, possibly due to implicit user assumptions.

User Preference for Situation Digest

At the end of the experiment, when we asked the subjects which view they preferred, a near-majority of the subjects responded with a preference for the digest view. Only 14 percent preferred the normal WEA view and the remaining 38 percent were either indifferent or unsure. The results were significant with a large effect size. See the last section of Appendix B5 for tabulated results pertaining to user preferences.

Overall Evaluation of Situation Digest

The longer and more complex an emergency situation is, the more likely the users are to make mistakes when interpreting the information provided by the regular WEA view. In these cases, the Situation Digest feature would have a significant impact on the users' understanding of the nature of the emergency and the action to take following an incoming alert. Users tend to interpret all incoming alert messages and updates as immediately actionable by nature. In terms of improved understanding of the immediacy of an emergency, this feature would thus make little difference.

Additionally, there is no downside to using Situation Digest for simple emergency cases when the information does not change much over time. The feature does not compromise any aspect of situational awareness in such scenarios. This result could facilitate the adoption of the feature if implemented in the real WEA service. Users would get used to the consolidated view in the simple cases. When a large disaster happens, the feature would already be in place and in use and the users would be accustomed to it.

Finally, over three times as many users preferred the digest view compared to the normal WEA view. Combining this positive overall impression with the feature's positive impact on situational awareness in complex scenarios without any apparent downsides leads us to conclude that having an alert creation and presentation scheme specifically catering to situational awareness is well worth considering in the future evolution of WEA.

4.12. Results: Overall Impressions About WEA Benefits

At the end of each of the two trials and three experiments, a final questionnaire was sent to find the overall perception of the study. In this final questionnaire, two questions were asked related to the WEA benefits:

1. Do you believe wireless emergency alerts are useful?
2. Do you believe wireless emergency alerts could save lives?

Analysis of the feedback reveals participants in each of the three experiments almost unanimously felt that WEA service was useful and that these alerts could save lives. In Experiments 1 and 3, respectively, all 20 (all 19 for question 1) and 37 subjects, who responded gave a positive response to both questions. In Experiment 2, of the 41 and 43 subjects who responded to question 1 and 2 respectively, 87 percent responded positively to the first question and 92 percent responded positively to the second question, again indicating a strong belief in WEA.

The pooled results therefore indicate a strong positive perception of WEA with around 94 percent of the subjects believing in the usefulness of WEA messages and 96 percent believing in their life-saving potential. High values of the Chi-square statistics suggest confidently rejecting the null hypotheses for both questions, with considerably large effect sizes of 1.29 and 1.33 as measured by the Phi statistic. Appendix B6 provides tabulated pooled results from all experiments.

4.13. Study Validity

This section discusses main threats to the validity of Trials 1 and 2, and the measures taken to alleviate these threats.

The results may have limited generalizability due to the particular demographics of the subject samples used. The subjects were predominantly technology savvy and comfortable using smart phones and their advanced capabilities. This population characteristic could have influenced the results by especially favoring the technology-intensive features tested. Certain differences between the findings of the two trials (specifically between Experiment 1 in Trial 1 and Experiment 2 in Trial 2) also suggest that some tested factors, such as use of TTS and alert timing, may be particularly sensitive to demographics. However, these same findings also happened to be inconclusive and did not result in any recommendations. Also in general, volunteers tend to be more motivated than the general population. The emergency services community in particular is known for its altruistic behavior. Therefore, there is a risk that the reactions and responses of the sample may not mimic those of the average, less-keen citizen. As an alleviating point, motivational factors apply equally to the compared groups in an internally randomized design, which is the case here.

Although the emergency scenarios were fictional, the majority of the alerts used in the trials were modeled on real alerts used in real emergencies. The only notable exceptions were the alerts used in the last scenario of Trial 2 during its final days, and this choice was deliberate to invoke suspension of disbelief, increase motivation and prevent too much focus on local details. Even in these cases, the form and language of the alerts mimicked examples we could find from various sources. The alerts' realism was evaluated in the final questionnaire and found to be reasonable by the subjects.

We did not identify any significant threats to construct, internal and statistical validity, with the exception of history threat (the possibility of the subjects' behaviors being influenced by uncontrolled external events), which likely mostly affected the factors with inconclusive findings.

At the end of each trial, a questionnaire was sent to all subjects to collect feedback on study design and execution. Questions asked were related to clarity of motivation, intrusiveness of trials, alert frequency, alert realism and effectiveness of kickoff event. Response rate was 38 percent for Experiment 1 (20 out of 52 subjects), 87 percent in Experiment 2 (47 out of 54 subjects) and 61 percent for Experiment 3 (37 out of 60 subjects). The responses indicate no red flags that could compromise validity.

4.14. Summary of Trial 1 and 2 Findings

Table 4.9 summarizes the findings for enhancement features and lists potential implications of each feature. Based on the significance of the findings, we rate the strength of evidence for each feature as *Strong*, *Moderate* or *Weak*. We also rate their potential impact as *Significant*, *Medium* and *Low* based on how many alert outcomes they positively influenced and to what extent. The three features with the strongest evidence incidentally have the highest potential impact on one or more outcomes evaluated. Features meriting this double rating fell into two categories and are considered *high-value*:

1. Low-cost and high-feasibility context-aware enhancements that rely on filtering on the phone and primarily impact alert relevance; and
2. A lower-feasibility and higher-cost enhancement that relies on a structured alert generation and presentation approach, and primarily impacts the recipients' sense making and confusion (which we represented as situational awareness).

It is important to note since none of the enhancements proposed require changes to the essential network architecture of WEA, they can be implemented within the existing architectural constraints. Some do require AOs to modify the structure and content of messages they send, however. These modifications are indicated under the *Implications* column of Table 4.9.

In the third column of Table 4.9, positively impacted outcomes for each tested feature are listed in order of strength, from strongest to weakest. Bold outcomes represent significant effects with acceptable effect sizes in both trials and in pooled results if they were tested in both trials or with medium or large effect sizes if they were tested only in one trial. The rest of the outcome measures were significant only in one trial or in the pooled results if they were tested in both trials.

The *Feasibility* column shows the researchers' assessment of how easy it would be to implement a given feature in future versions of the WEA service depending on the changes required to the alert creation and delivery pathway. This naturally inversely correlates with the feature's cost. We have demonstrated the feasibility of all the listed features in an emulation environment by implementing both server-side (alert origination) and client-side (phone functionality) proof of concepts in the WEA emulation testbed.

Table 4.9. Summary of Findings and Implications for Tested Enhancement Features

WEA enhancement feature	Evidence in favor / Potential for improving WEA	Positively impacted outcome constructs	Negatively impacted outcome constructs	Feasibility	Implications
Geo-targeting: fine-grained, precise targeting on the phone using recipient's location	Strong / Significant	1. Relevance 2. Actionability 3. Adequacy	Hindsight relevance	High	Requires embedding geo-target into alert at origin and enabling of location services on phone for geo-filtering. Straightforward to implement on phone with no user interface (UI) changes. Invisible to AO and recipient. Supported by polygon/text compression, moderate increase in longer message length.
Situation Digest: consolidated view showing up-to-date summary for each active emergency	Strong / Significant	1. Situational awareness		Medium	Requires new alert creation process/tools and use of CAP metadata. Metadata encoded and bundled with alert content for transmission and unbundled or decoded on phone. Visible to both AO and recipient. Changes phone functionality, including UI. No changes to the WEA network architecture are necessary.
Location History: extended delivery to citizens who are likely to visit the alert area based on past movement patterns	Strong / Significant	1. **Relevance** 2. Usefulness 3. Actionability 4. Adequacy	Hindsight relevance	High	Used with geo-targeting. Straightforward implementation on phone possible with no UI changes. Can be invisible to AO and recipient. Effectiveness depends on prediction algorithm.

WEA enhancement feature	**Evidence in favor / Potential for improving WEA**	**Positively impacted outcome constructs**	**Negatively impacted outcome constructs**	**Feasibility**	**Implications**
High-Information Map: alerts augmented with maps showing recipient location and geo-target	Moderate / Medium	1.Usefulness 2. Relevance 3. Adequacy		High	Used with geo-targeting. Implementation on phone is straightforward, but requires new UI. Pre-cached maps on phones alleviate network congestion concerns.
External Link: inclusion of an external link in the alert content	Weak / Low	1.**Usefulness** 2. Adequacy		High	Requires only minor changes to phone app to make links clickable. Policy change is necessary to allow (restricted) links in alerts. Network congestion may be a current concern, but is likely to be solved by advances in cellular technology.
Location Prediction: extended delivery to citizens who are likely to visit the alert area based on proximity and motion	Weak / Low	1.Usefulness	Hindsight relevance	High to Medi um	Used with geo-targeting. Implementation on phone possible with no UI changes. Prediction algorithm can be complex. Can be invisible to AO and recipient. Effectiveness heavily depends on prediction algorithm.
Long Message: alert messages longer than 90 characters and carrying more information	Weak / High	1.Actionability 2.Annoyance 3. Milling behavior		High	No changes in alert creation side beyond AO training and modifications to existing tools to enable longer message construction. Requires policy change. Supporting extra capacity implies upgrades and changes to alert transmission infrastructure and protocols. Cooperation from wireless carriers is required. Potential to support geo-targeting-related features increases importance.
Text-to-Speech: audio notification of incoming alert using spoken content of alert	Weak / Unknown	1. Response Rate 2.Understanding 3.Actionability	Response time (short-term)	High	No changes in alert creation beyond avoidance in alert text of content not easily converted to speech. Simple changes to phone app.

Alert characteristics tested, namely alert timing (day-time vs. night-time alerts) and update alerts (alerts announcing a new emergency vs. alerts updating information sent in a previous alerts), are not included in the summary table since we do not have any conclusive findings about them. We will nevertheless briefly discuss their implications further below.

4.15. Recommendations: Features with a Significant Potential for Improving WEA

Among all the tested enhancement features, Geo-targeting, Situation Digest and Location History have the highest potential to improve the WEA service.

Geo-targeting. Fine-grained, precise geo-targeting based on filtering on the phone is the most significant finding of the trials. In particular, it improved the alert messages' relevance to the recipients markedly.

This feature is straightforward to implement using modern smart phones' location services, as was demonstrated in the Android WEA+ app. No user interface changes are required on the phone's built-in app that intercepts the WEA messages: the feature is completely transparent to the users. However, it requires the representation of the alert region, the geo-target specified by a polygon with geographical coordinates as vertices, to be attached to the alert message for processing by the recipient's phone. This would be the only change required on the alert origination end by minor modifications to the gateway functionality, but the change would be transparent to the AOs themselves.

Using computationally cheap and efficient compression and decompression techniques, we can significantly reduce the size of both the polygon and the alert text, allowing even the current 90-character limit to be sufficient to carry both the main payload and the geo-target in most situations. A moderately increased message length limit would make this enhancement extremely feasible in future WEA implementations.

Recommendation 1. Based on its high feasibility and potential impact, we strongly recommend that fine-grained, precise geo-targeting based on filtering on the phone be given serious consideration in the future WEA service.

Situation Digest. The next significant feature in terms of its potential to improve WEA is the use of a consolidated view to facilitate sense making, especially under complex emergencies involving multiple incidents. Such a view presents the users with the latest information about an emergency scenario. The Situation Digest feature implemented and tested this novel concept in the second trial.

In the standard WEA service, alerts are treated as a flat, unstructured stream of independent messages. Situation Digest is actually more than a simple feature; unlike in standard WEA service, Situation Digest also requires AOs to think of alerts and alert streams relating to both separate and inter-related emergencies in a structured way using the full capability of the CAP. Our experiments showed this concept can significantly improve the recipient's sense making, or situational awareness, in a rapidly evolving emergency situation involving multiple and interrelated, incidents as might typically happen in a large disaster.

We measured situational awareness using the recipients' time-dependent, aggregate understanding of the type, recommended action and immediacy of an underlying emergency scenario. Overall, there was a marked improvement in situational awareness with the digest view, compared to the single-stream normal WEA view. The complexity of the emergency scenario presented amplified the improvement, and no compromise was observed in simple scenarios.

It is not as straightforward to implement and operationalize this concept as other features, although no changes to the standard WEA network architecture are necessary. The concept leads to a new way of thinking about emergency alerting. It requires metadata defined in CAP to give structure to alert messages, such as explicitly specifying an incident identifier, and other fields representing different alert attributes, such as type of incident, action to take and urgency. This in turn affects the alert creation workflow, necessitating new tools with new capabilities at the originating end. The implementation, as proposed here, also requires the metadata to be efficiently encoded and bundled with the alert payload, just like the geo-target (indeed the geo-target is just one component of CAP). In our recommended implementation (Iannucci et al. 2015), CAP messages carrying the alert content and metadata are encoded as a binary blob, and the gateway transporting messages from outside the communications network to inside the network needs to be able to create a SMSCB message out of this blob and set the language byte as binary. This change may involve cooperation of participating wireless carriers.

Phone-side changes include new functions for digesting the active alert streams, a well-designed user interface and a signaling protocol, similar to the one implemented in the WEA+ testbed, for updating the user view automatically. On the receiving end, the message is unbundled, de-coded and the metadata is analyzed to generate a digest view. All of this processing is done locally on the by the phone when an incoming alert is intercepted. The user interface of the built-in WEA app on the phone must present this information to the user in a clear and organized way; hence, the success of the feature is dependent on human factors as well. Our presentation strategy for the Situation Digest feature was simple; it was not designed with the help of a user experience expert, but the feature was still valuable.

Therefore, the changes necessary to implement a similar digest-based notion in a real context would impact many components of the alert creation and delivery pathway. The changes would also require policy revisions as well as involvement from several stakeholders.

Recommendation 2. Accommodating situational awareness through a digest view is a high-impact, high-value concept that future WEA service should consider implementing. Such an implementation could be complicated by public-facing and AO-facing implications involving multiple stakeholders. Given the evidence, we believe the benefits would far outweigh the costs, however.

Location History. Location-history-based filtering combines geo-targeting with interest-based targeting. Interest-based targeting allows users who have a sufficient interest in an alert's geographical area to receive messages even when they are outside that area. Filtering based on location history achieves this by comparing a continually updated record of recently visited locations on the user's phone to the geo-target of an incoming alert to infer the user's interest. We implemented and tested this feature only in the second trial. In geo-targeted alerts, the use of location history to make alert delivery decision significantly improved the alert messages' relevance to the recipients compared to unfiltered alerts. To a lesser degree, the perceived usefulness, actionability and adequacy of the alerts were also positively and significantly affected.

The feasibility of location-history-based filtering is subject to similar constraints to precise geo-targeting. It does not require any significant changes on the alert origination side besides the inclusion of the geo-target into the alert. The filtering algorithm can be implemented on the phone locally with no changes to the user interface. However, effectiveness in actual use of the feature may depend on the sophistication of the algorithm itself and whether the users are allowed to tune it to reflect their own preferences. In the trial, we used a simple scheme that proved sufficient.

Recommendation 3. Using location history of users to filter incoming alert messages is another way of leveraging advanced capabilities of modern smart phones. Simple automatic filtering schemes enabling the delivery of alert messages to people who have an interest in an alert's geographic area, but happen to be outside it, are relatively easy to implement without any impact on alert creation and transmission beyond what is required for precise geo-targeting. Such schemes are highly likely to improve alert outcomes and alleviate opt-out. If necessary, they can be tuned and customized by the users to fit their specific needs. We thus strongly recommend that future WEA service offer this capability as an adjunct to geo-target-based filtering.

4.16. Recommendations: Features with Moderate Positive Effect Worthy of Consideration

The next group of enhancement features in terms of importance and potential are Long Message, High-Information Map and Location Prediction. Among these, we consider Long Message to be an enabler; we include it in this second group not based on direct evidence of systematic, positive impact on alert outcomes, but because of its potential, indirect impact in supporting other, high-value features.

Long Message. Increasing the current 90-character length limit of WEA messages has been under consideration for some time. We consider the ability to send longer messages to be an enabler for other features rather a feature of high importance in itself. We did not find persistent significant differences in alert outcomes between long and short messages, although in certain cases a minor effect on actionability, annoyance and milling behavior was observed. These findings contrast with results of some recent studies, for example Hamilton et al. (2014), that predict an independent, systematic improvement in alert outcomes with messages considerably longer than 90 characters.

From our perspective, the most practical use of long messages would be their ability to carry extra metadata required for filtering-based enhancements such as the precise Geo-targeting and Location History features, as well for structured alerting schemes catering to situational awareness, such as the Situation Digest feature. We have demonstrated that even with the message-length limit, in most cases, we can accommodate extra metadata with the use of clever, computationally cheap compression and encoding algorithms. A moderate increase in message length, say to 180 characters, would create sufficient additional capacity to cover the remaining cases.

Recommendation 4. We recommend that future WEA service consider allowing longer alerts to create extra capacity for attaching or bundling metadata with alert payloads. A moderate increase in message length limit would be an important enabler for future WEA innovation driven by features relying on metadata.

High-Information Map. Perceived as a highly useful and desirable visual adjunct to geo-targeting, a high-information map communicates to an alert's recipient easily digestible information about the alert region and the user's proximity to it. Alerts with high-information maps leverage modern smart phone capabilities in a straightforward way. Provided that network congestion due to on-demand fetching of maps from Internet resources is not a concern, high-information maps are very feasible to incorporate to future WEA service. It requires no changes on the alert origination side beyond what is needed in precise geo-targeting. Network congestion concerns are readily addressable by pre-cached maps, which are expected to be ubiquitous in future generations of smart phones. Adding a high-information map to a geo-targeted alert feature is fairly easy to implement, although it requires a new recipient-facing user interface on the phone, which necessitates some changes to the client functionality interpreting and displaying incoming alerts.

Recommendation 5. Since high-information maps are highly desired by users, highly feasible to implement, and have the potential to improve alerts' relevance and information content, they should be given serious consideration in future WEA service. Concerns for potential network congestion could be addressed by making the feature selectively available on client devices based on device capabilities (e.g., availability of pre-cached maps) and real-time connectivity (e.g., Wi-Fi vs. cellular data), but we expect these concerns to diminish over time with advances in LTE technology and smart phone functions.

4.17. Features and Characteristics with Small or No Demonstrable Effect

The last group includes features and characteristics tested, but not yielding salient enough results to warrant a clear recommendation. Weak or inconclusive results were likely due to a combination of the relatively small subsamples used in their testing and various confounding factors threatening the validity of their testing. These threats were discussed in the study validity section. For this reason, we think that most of features and characteristics in this group deserve further investigation with new and improved study designs conducted with larger subsamples.

External Link. Clickable external links, in the form of URLs or social media hash tags embedded in alert messages, are perceived to be highly useful and easy to implement. Their hypothesized positive impact on other central alert outcomes could not be confirmed, however. External links appeared to improve the information content of alert messages, but the observed effect was small and not systematic. Given the underlying concern about network congestion, we consider this feature to be a feature worthy of exploration and future study, but cannot make a clear recommendation at this time. We do not expect network congestion to remain a real problem (although it may persist as a perceived problem in the near future), since we believe it will be easy to address with impending advances in cellular technology.

Location Prediction. Filtering based on predicting a user's location from the user's direction of movement, speed and proximity to the alert area is another client-side adjunct to precise geo-targeting. The implementation of this feature on the phone may require sophisticated algorithms to make it robust enough. In the second trial, we used a simple algorithm which failed to yield significant results despite the favorable rating of this feature by users. We are not certain whether the lack of evidence concerning the remaining alert outcomes was due to the shortcomings of the implementation, underlying concept itself or small sample size.

Text-to-Speech. TTS is an easy-to-implement, attention-grabbing feature that augments alert notification with spoken text. We had mixed results with TTS. Its ability to attract the attention of the users was validated by an increased response rate; however, findings regarding other alert outcomes were either weak or contradictory.

Update Alert. By comparing initial alerts announcing a new emergency to update alerts modifying the information included in a previously sent alert, we observed that initial alerts outperformed update alerts in prompting information-seeking behavior, and conversely update alerts improved information content. These results do not lead to a specific actionable recommendation, but it motivated the subsequent treatment of update alerts in a more structured way than as in the standard WEA service in order to tackle situational awareness.

Alert Timing. We compared alerts issued during the day to alerts issued during the evening and at night to find out if alert outcomes were affected by alert timing. However, either there were no significant differences or we were not able to isolate this alert characteristic well enough. Thus we are not able to draw further conclusions or make actionable recommendations with regard to alert timing.

4.18. Geo-target Compression Tests and Results

The AORS interviews and survey, as well as other sources (DHS 2013; McGregor et al. 2014; SEI 2014), indicate that two of the most important improvement desires for the WEA service are for longer WEA messages and for the ability to include the geo-target and other potentially useful information directly in the message. Given the current 90-character message limit and the expectation that even if future WEA service supports longer messages, a message length limit will still be in place, we launched a key subproject to explore the feasibility of substantial compression of message text and of embedded geo-targeting polygons.

We had access to a corpus of approximately 11,000 WEA messages sent out by the NWS (Gerber 2014). This corpus contained 11,368 polygons collected from the NWS online portal from 2012 to 2014. A careful examination of the nature and structure of the message texts and the polygons used inspired a number of new heuristic compression schemes, as well as an examination of conventional compression schemes. We experimented with these schemes to shorten the alert text and polygon representation separately.

The polygons in the NWS corpus ranged from 4-24 points, with a size ranging from 49-331 characters. Traditional polygon compression techniques that manage large numbers of connected polygons do not work well on these short polygons. Therefore, new techniques were devised. The NWS polygons tended to be relatively small, with most coordinates near each other and with coordinate precision limited to two decimal places. We achieved substantial compression by converting all coordinates to deltas from neighboring coordinates or a known origin and converting the numeric base from decimal using 10 characters (0-9) into a higher base using 62-90 characters (0-9, A-Z, a-z, …). The best techniques applied several heuristics in combination to perform an initial compression, and then other algorithmic techniques, including higher base encoding on top of the initial compression. Further, these methods were respectful of computation and storage constraints typical of cell phones. Two of the best techniques included a “bignum” quadratic combination of integer coordinates and a variable length encoding which took advantage of a strongly skewed polygon coordinate distribution. Both techniques applied to one of the two delta representations of polygons are on average able to reduce the size of the input by some 80 percent.

Using these techniques, we were able to compress NWS polygons used to target WEA weather alerts to between 9.7 percent and 23.6 percent of their original length, depending on specific characteristics of individual polygons. This compression range implies that we could reduce original polygons from 43-331 characters to 8-55 characters, to fit well within even the current 90-character limit.

Figure 4.8 compares the original message lengths (marked with O) with the lengths yielded by the various compression methods. Message length is shown on the vertical axis and the number of polygon points is shown on the horizontal axis.

We also examined standard techniques such as Golomb (Golomb 1996) and Lempel-Ziv-Welch (LZW) (Ziv and Lempel 1977). Two of the best methods turned out to be our heuristic approach using base-62 bignum arithmetic (with a mean compression of 23.90 percent and standard deviation of 10.36 percent) and another using a base-62 LZW-like method with a static dictionary (with a mean compression of 24.11 percent and standard deviation of 12.32 percent). Even Microsoft’s Point Compression Algorithm (Microsoft 2015) does not perform as well as our best methods, probably due to its lack of the concatenation step that combines each polygon vertex in our method. As can be seen, all methods are essentially linear in the number of points, and thus our methods are scalable and compress larger polygons extremely well.

Full details on the devised polygon compression methods and their evaluation can be found in a companion report at the 2015 American Meteorological Society (AMS) Weather Warning conference (Jauhri, Griss and Erdogmus 2015).

The next step in polygon compression involves experimenting with smaller polygons with higher precision coordinates, more typical of a city block or street. To identify the best heuristics and algorithms, we would need a similar corpus of typical polygons. We expect that the same type of techniques will work here as well, though some compression parameters might be different.

To compress message texts, we developed a different set of preliminary techniques and the results were equally encouraging. In the NWS corpus, the 11,000 WEA messages had a distinct structure and repeated use of a small set of common words. They range in length from 68-90 characters. Examples are:

- "Flash Flood Warning this area till 2:00 AM CST. Avoid flood areas. Check local media. -NWS"
- "Tornado Warning in this area till 11:30 AM CST. Take shelter now. Check local media. -NWS"
- "Dust Storm Warning in this area till 12:00 AM CDT. Avoid travel. Check local media. -NWS"
- "'Ice Storm Warning this area till 6:00 AM CST. Prepare. Avoid Travel. Check media. -NWS"

The observed repetition and structural similarity mean a compression scheme using a small pre-computed dictionary of common words, multi-word phrases and strings found in WEA messages indexed by one or two characters will yield substantial compression. We explored several alternative ways of generating the dictionary and encoding the output. The simplest technique was to consider single or multi-word phrases like "Flood Warning this area," "Check local media," "-NWS," etc., in the WEA corpus and rank them based on number of occurrences. A manually constructed dictionary with 120 word and phrase entries compresses an original message of 68-90 characters to 18-46 characters. An automatically built dictionary of the 147 unique words in the corpus compresses the original NWS WEA messages to 13-32 characters. If we allow both single and double words, the dictionary then grows to 493 entries, with messages compressing to 8-17 characters, some 12 -19 percent of their original length. Going to three word phrases lengthens the dictionary substantially, but provides only minor compression improvements.

Such techniques would not efficiently utilize all phrases at compression since redundant phrases like "Flood Warning this area" and "Warning this area," which have same number of occurrences in the corpus, will be present in the dictionary. To avoid such redundant entries in the dictionary, we considered a greedy approach, which only keeps the largest phrase in the dictionary if multiple overlapping phrases exist. For example, the greedy approach will only keep "Flood Warning this area" in the dictionary and avoid phrases like "Warning this area" or "this area." This optimization reduces the dictionary size from 493 to 349 entries, with a slight increase in the compressed message lengths to 8-22 characters. There are 36 words that only appear once in the 11,000 messages, and if we were to only use repeated words or multi-word phrases, we have a small numbers of outlier messages that do not compress very well.

In conclusion, using the best of the polygon and text compression methods, we could readily fit a compressed message and essentially all of the polygons into the current allowed 90 characters, with a relatively small dictionary to be stored on the phone. A larger character limit would easily allow longer messages and higher resolution polygons with more vertices, still leaving space for other types of metadata to support various future WEA enhancements.

The results of the compression work are promising and could have a large impact on the evolution of the WEA service. They were shared at a Communications Security, Reliability and Interoperability Council's (CSRIC) meeting and the 2015 AMS Weather Warning Conference. Plans are under way to discuss the findings at a future Federal Communications Commission advisory meeting.

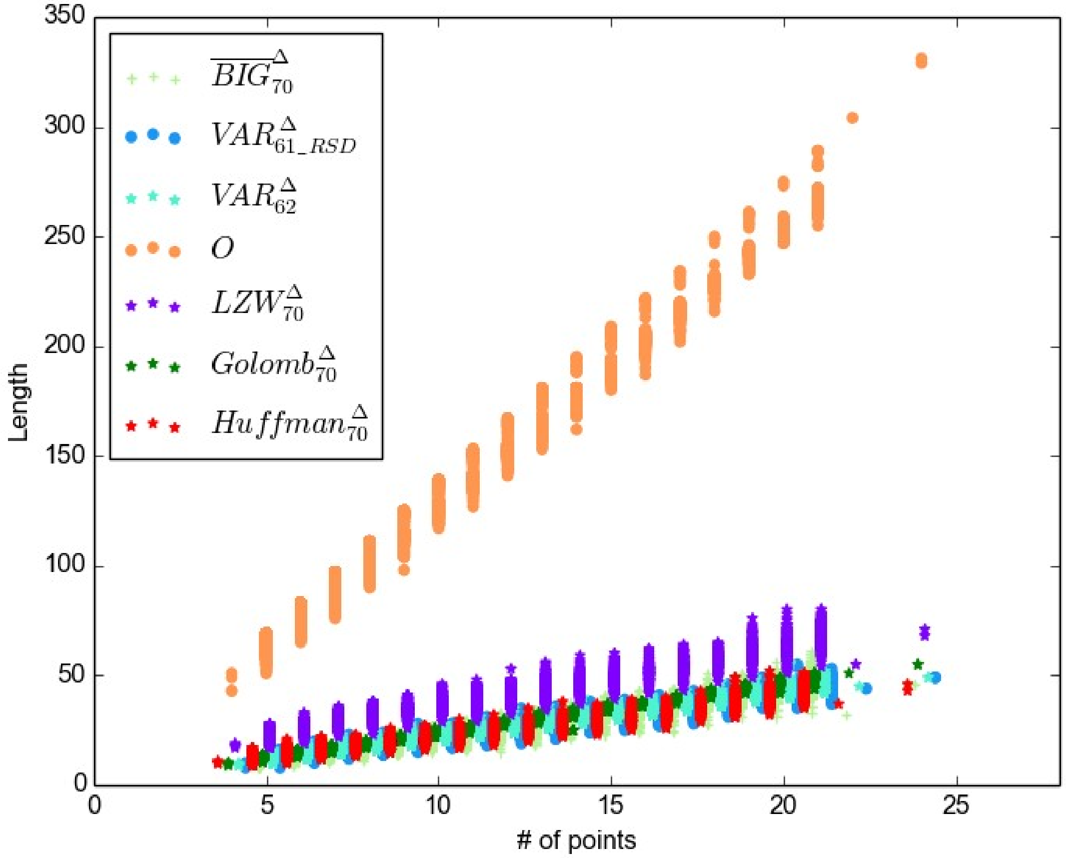


Figure 4.8. WEA NWS Corpus Geo-Target Polygon Compression Results Using Different Techniques

5. Overall Conclusions and Recommendations

Our studies have led to insights about the perceived current value of WEA and the potential future value that could come from specific improvements. In this section, we present the five major themes that emerged as the best overlap of "what's valuable" with "what's possible." For each, we describe the theme in the broad context of emergency alerting, offer supporting evidence from the three studies, and outline the potential system-level costs and complexities of potential improvements. For each theme we conclude with a summary recommendation. We prioritize the recommendation on a scale from 1 (highest priority) to 5 (lowest priority). This priority is not a mathematical aggregate of all factors. It is, rather, our considered opinion based on the aggregated supporting evidence as we emerge from 18 months of WEA surveys, experiments and analysis. As such, there may be some debate between priority-three and priority-four recommendations as to which is the more important, but there is little doubt in our minds that priority-four should be given clear preference over priority-one in terms of future focus.

It is important to note that the AO interviews, AO survey and PUT yielded a wide variety of significant insights, and the list presented here is by no means exhaustive. But it is our collective view that these are the five meriting serious consideration:

1. Deep integration of location-based context materially improves WEA's value.
2. WEA must transition from a focus on *alerting* to a focus on *awareness.*
3. WEA must be well-positioned as a peer in the social communication pantheon.
4. Rich-media integration into WEA is a question of *how*, not *if*.
5. Education, testing and measurement are necessary for WEA's success.

In Section 1.1, as part of our starting assumptions, we introduced the *opt-out problem* (the propensity of the public to disable the WEA capability on their phones or stop paying attention, if they find incoming alerts irrelevant), the short-message fallacy (the belief that the WEA service must be limited to short, unstructured text messages), and the location proxy fallacy (the belief that current location is the single most important attribute determining a recipient's interest in an alert). These five insights help attack the opt-out problem as well as debunk the short-message and location-proxy fallacies. Table 5.1 summarizes the findings and recommendations related to these insights.

5.1. Deep Integration of Location-based Context Materially Improves WEA's Value

One of the main motivators for this work was to prevent or reduce *opt-out* — the process by which users, frustrated with too many irrelevant WEA messages, switch the functionality off in their smart phones or ignore incoming messages. To that end, we explored a line of research aimed at evaluating both the feasibility of smart-phone-based geo-targeting and its value (or not) as measured through controlled experiments during the PUT phase of the research.

We found strong support for the concept in both the AO interviews and the AO survey. AOs in general, especially those from city- and county-level jurisdictions, thought that increasing geo-targeting resolution to permit more precisely targeted alerts and alerts targetable to a small area would significantly increase the effectiveness of the WEA service. We demonstrated feasibility of smart phone-based geo-targeting, and through the controlled experiments of the PUT phase, we uncovered a high level of acceptance of and value for the deep integration of location-based context. In particular, phone-based, fine-grained geo-targeting significantly improved the perceived relevance of alerts to the recipients, which was the strongest finding of the PUT phase.

Service-level implications: Smart-phone-based geo-targeting represents a departure in approach from classic WEA. Today, the interpretation of targeting information is a function of the cellular network. We propose to curtail such source filtering in favor of over-messaging with smart phone interpretation: sending specially-coded alerts to a broader-than-previous population, relying on logic to be built into smart phones to interpret the transmitted geo-target in the individual subscriber's context and issuing the alert (or not) accordingly. This approach is feasible because the interpretation software would consume very little resource from the phone, yet could provide a powerful mechanism for targeting (alerting the user not just based on his/her current location being in the target area but also of his/her location history — or that of selected family, friends and associates — being in the target area).

We have further demonstrated the means by which targeting polygon information can be efficiently coded and embedded in a WEA message using an efficient, lossless compression scheme (Jauhri, Griss and Erdogmus 2015). To evaluate and characterize the effectiveness of this scheme, we applied it to polygons and CMAM[8] free text fields of CAP messages from approximately 11,000 historical WEA alerts provided by the NWS. The original polygons from each alert consumed 43 to 331 characters. After compression, they were reduced to 8 to 61 characters, representing compression ratios of 9.7 percent to 23.6 percent. Likewise, the original CMAM text strings were 68 to 90 characters in length. Compression reduced them to 8 to 22, with corresponding ratios of 11.8 percent to 26.5 percent.

It is our recommendation to explore a separation of WEA into a transport component with minimal message interpretation by the network and an interpretation component residing solely and securely on a subscriber's smart phone. As smart phones quickly become ubiquitous, this approach is increasingly feasible and increasingly valuable. We note that it may well be possible to code such messages in a way to not trigger spurious over-targeted alerts in older "feature phones" — meaning that both old WEA and smart phone WEA can coexist on a network without mutual interference.

5.2. WEA Must Transition from a Focus on *Alerting* to a Focus on *Awareness*

As the PUT has shown, simple text alerts are satisfactory for creating an understanding of a simple situation and the actions to be taken. But if the situation becomes multi-dimensional (such as has occurred in an earthquake resulting in multiple fires, traffic snarls, water shortages and other issues) and/or evolves over time (such as when a prior instruction has to be modified or rescinded), sequences of text alerts with inter-related messages can lead to confusion and incorrect actions with potentially serious consequences. Alerting alone is inadequate. The stitching-together of multiple messages to create understanding or awareness is essential in such complex cases. Should this burden rest on the shoulders of the general population, or can WEA be improved to provide this message digesting function? Our research shows support from AOs and from the PUT for a fundamental change in WEA's focus from issuing alerts to creating awareness, pointing to abilities to digest complex situational information and to support the needs of complex and evolving situations.

Service-level implications: We have demonstrated a system for enriching WEA that can encode, transport and interpret on-phone multiple inter-related CAP messages as an overlay on the existing service. This enrichment allows us to convert WEA from an *alerting* service to a *situational awareness* service with measurable improvement in understandability.

Through the PUT and our related research, we have demonstrated both the value of this change and the means by which it can be retrofitted into today's WEA at minimal per-message cost and incremental system complexity. In fact, our approach based on CAP *preserves* information already in the alert origination system rather than discarding it as is done with today's WEA.

---

[8] Commercial Mobile Alert Message

We further observe that improvements in the underlying transport technologies (e.g., MBMS) will only serve to further enhance an awareness-based WEA.

5.3. WEA Must be Well-positioned as a Peer in the Social Communication Pantheon

In our research, we uncovered strong evidence that social media have become a key real-time news source and, in particular, geographically-local groups and/or services (e.g., NextDoor) are being used regularly by local police for two-way communication with residents. The question is well posed: if social media have this level of acceptance, what role should WEA play?

In our interviews, we revealed a strong sense among AOs who are well-versed in social networks that WEA serves a critical role as an *alarm bell*. But the inherent limitations of WEA further suggest that it cannot and will not be the only source to which the public will turn. Interestingly, as we studied actual incidents and how each appeared in WEA and social media, we produced at least weak evidence that there is little/no *footprint* of WEA in related social media — suggesting that discussions of events emerged on social media with no dependence on WEA as the alarm bell. This may be due to insufficient WEA presence (AOs may not be trained, may not have chosen to use WEA, did not use it effectively or official WEA alerts come later than the start of social media buzz), but the research suggests that many people are getting their first alerts from something other than WEA. This calls many aspects of the practice of WEA into question and motivates additional study. Simply seeking to measure time-to- awareness and source-of-awareness across a population when messages are delivered simultaneously via WEA and social media would give some basic understanding of where WEA stands today in the presence of social media.

Anecdotally, we found evidence of a rapid ramp-up in use of social media at the City of Sunnyvale who has transitioned in just a few years from little use of social networks to regular, and very successful use (with a very high level of opt-in from the residents). The speed at which the transition took place is especially noteworthy and indicates to us the public's tacit acceptance of social networks as a valued source of information overlapping the scope of WEA. As a practical matter, in times of emergency, it is human behavior to rely on known and trusted communication tools. If services such as NextDoor are known and trusted, and WEA is infrequently exercised, which one will AOs reach for in a time of stress?

Service-level implications: It is our opinion that WEA as a service must be re-formulated as a complement to existing and future social networks and that the alert origination processes should be augmented to integrate both traditional IPAWS targets with social media in a way to preserves consistent identities (e.g., the NWS is clearly and consistently branded and represented in WEA, Facebook, Twitter and NextDoor) and messaging so as to avoid confusion and effect the fastest and broadest-possible alerting and follow-up.

5.4. Education, Testing and Measurement are Necessary for WEA's Success

Education: Our interviews and surveys revealed the need for education of both AOs themselves and the general public. This feedback correlates with findings of recent previous studies (SEI 2014; Hamilton et al. 2014). We were surprised to learn how few AOs in our study had actually used WEA. While a small number were approved to issue messages or were in the process of getting approval, many lacked the need or the ability to issue WEA messages. Several AOs specifically identified the need for better education of both AOs and the public to improve the understanding, acceptance and use of the service.

We conclude that systematic education of AOs is well motivated to:

- Develop familiarity with alert origination tools and their relationship to downstream messaging systems;

- Instill confidence as to how the system performs; and
- Practice the art of condensing complex situations to 90-character messages.

Testing and Measurement: AOs further highlighted the need for regular systematic testing of WEA. Not unlike Civil Defense, the EBS and the EAS, AOs believe that to be effective in times of need, the WEA service needs periodic testing involving large populations. Our experience strongly suggests the development of a more general-purpose environment in which alert AOs, government agencies and emergency organizations can safely pre-test WEA concepts, features or enhancements with volunteers before these improvements are deployed can be highly beneficial. Such pre-testing would precede and complement wider-scale regular public testing.

Service-level implications: Most of the benefits of *education* are available with no changes to WEA as implemented today. For *testing and measurement*, the testbed created for this study enabled such prototyping, demonstration and evaluation of potential WEA improvements. Our alert origination, dissemination and on-phone-filtering and digesting tools and technologies, together with our CROSSMobile software-defined network, make for a powerful research and development platform. The insights gained through its use were invaluable. An opportunity now presents itself to expand our testbed to be broadly accessible and larger-scale in support of growing and maturing the WEA service.

Beyond traditional one-way alerting modalities, WEA as a service has the potential of offering a way to feed back to the AOs either general or specific information about the receipt of alerts, how many read them, associated actions and geographic clustering of responses. Using smart phone capabilities for geo-targeting and/or situational awareness opens the door to such closed-loop operation. Feedback data can be selectively returned (statistical sampling orchestrated by smart phones themselves), offering actionable insights to AOs without creating significant network congestion or privacy concerns. Experience with such feedback systems on a regular basis would materially improve our collective abilities to understand and improve WEA as a service. It is our opinion that it is possible to design such a system to be both accurate and privacy preserving. Further study is indicated.

5.5. Rich-media Integration into WEA is a Question of *How*, not I*f*

Our research, like that of others, provides support for the integration of rich media into WEA. We believe this arises from the fundamentals of (a) widespread use of smart phones and (b) information authoring and presentation being HTML[9]-based. Apps on both iOS and Android platforms easily integrate HTML-rendering views (portions of the screen real-estate), giving app developers the ability to bring web content into apps. Authoring tools for the web itself, and all the web sites from news agencies and others, stand as proof that HTML is here to stay as a proven and well-exercised way to represent the world's information. Popular email readers create and display mail that includes HTML content. It is only natural to expect WEA to rise to both the potential and the expectation of rich media messaging.

Service-level implications: Objections to rich media WEA are losing credibility with the natural advances in cellular network architecture. The classic arguments about network overload are much more easily overcome with technical capabilities such as the Multimedia Broadcast / Multicast Service (MBMS, or eMBMS in LTE), a component of LTE since Release 9 (3GPP 2014) (offering the potential to deliver rich content to a large population in a one-to-many fashion) and cell-based content caching (Oswal and Iyer 2010). And combinations of the two (merging common requests from multiple users aimed at a common URL and servicing these through LTE broadcast out of an eNodeB-based content repository) are well worth investigation. MBMS contemplates using SMSCB (the vehicle for current

[9] Hypertext Markup Language

WEA) as an advertising mechanism for MBMS streams (3GPP 2015) — *thus facilitating both the purpose of the alarm bell and the purpose of bandwidth-efficient rich media delivery of alerts*. In our opinion, the use of MBMS broadcast as a standard for enriched WEA should be pursued with high priority.

Table 5.1. Major Findings and Recommendations

Major Themes	AO Interviews	AO Survey	PUT	Comments	Priority (1: high, 5: low)
Deep integration of location-based context materially improves WEA's value	**Strongly Supporte d**	**Strongly Supported**	**Strongly Supported**	**Cellular location services are pervasive and driven by commercial interests; built-in maps will emerge**	**1**
WEA must transition from a focus on *alerting* to a focus on *awareness*	**Moderately Supported**	**Supported**	**Strongly Supported**	**Value is demonstrable; retrofit to current WEA possible; even more compelling with rich- media integration**	**2**
WEA must be well-positioned as a peer in the social communication pantheon	**Strongly Supporte d**	**Strongly Supported**	**N/A**	**Opportunity to re-think alert origination; integrate alert message targeting with social media; both "pull" and "push" modalities needed**	**3**
Education, testing and measurement are necessary for WEA's success	**Supported**	**Strongly Supported**	**N/A**	**Education is necessary but not sufficient: deep problems remain in both technology and practice of WEA that education alone will not fix**	**4**
Rich-media integration into WEA is a question of *how*, not *if*	**Moderately Supported**	**Qualified Support**	**Neutral**	**MBMS broadcast, cell-level content caching are emerging; HTML-based authoring and rendering tools are pervasive**	**5**

6. FUTURE RESEARCH OPPORTUNITIES

While our research addressed a number of issues of central importance to WEA effectiveness and adoption and provided evidence-based answers, many open questions remain and several improvements are possible. These questions and improvements are discussed here as possible focus of future work and are offered in addition to the recommendations under Service-level Implications in Section 5.5.

Improvements to and subsequent re-testing of existing enhancement features. Some of our findings were sensitive to the specific schemes and algorithms used in the implementation of the features. The algorithms used for filtering alerts based on location history and location prediction are examples. In particular, the location prediction algorithm was a first approximation and can be made more robust, for example by taking advantage of activity recognition and machine learning techniques. The Situation Digest feature we developed to address situational awareness is another example: it depended on a simple user interface. While the results of the experiment were very positive, it is our sense that further improvements are possible by focusing on the psychological aspects of how such information is presented and assimilated, particularly in a stressful context. This in turn may improve measured outcomes. Map visualizations could be extended with additional information, for example showing the recipient's distance from the alert zone, estimating time of entry into the affected area, or calculating and highlighting the nearest route for evacuation. Performance improvements in some features would also be desirable; in particular, location histories could be stored more efficiently to save phone memory and speed up filtering decision for location-history based filtering.

Implementation and testing of new context-aware enhancements. We only barely scratched the surface in exploiting contextual information available on modern smart phones. Smart phones equipped with a multiplicity of sensors are now able recognize their users' physical activity in real time. Indeed, in Experiment 3, we are able to infer and categorize user activity (whether currently moving in a motor vehicle, cycling, running, walking or stationary), but did not take advantage of this information in any of the features tested. The phone may use activity information to make a balanced delivery decision to avoid distracting the recipient in dangerous situations, for instance, when one is driving at a high speed on a motorway, or changing the notification mode, for example by switching to TTS. Other contextual information includes alert timing, users' explicit preferences and interests (through user configuration and customization) and users' learned preferences and interests based on phone-based sources such as social networks, call histories and chat histories. These latter sources may raise additional privacy concerns even though the information used does not need to leave the phone. Beyond location-based ones, we have not yet explored any of these options.

Two AOs in our study reported the need to address multi-language and translation issues for any type of an emergency message to be sent out. Modern software engineering practice segregates the logic of applications from the localized text strings. Integrating user- and context-specific localization builds on our work in context-aware filtering. Using the proposed situational-awareness CAP transport mechanism opens the door to richer encoding of alerts including multiple languages, with on-phone software choosing the appropriate language(s) for presentation.

Further testing of implemented enhancements with expanded populations and in longer-duration studies. Replication of experiments with wider and more diverse populations more representative of the general public would strengthen the evidence gathered, improve the generalizability of the results and help re-evaluate inconclusive, counter-intuitive, surprising and inconsistent findings. Longer-duration experiments would allow optimal tuning of the parameters of certain features tested. For example, location-history-based filtering showed good promise despite even though it relied on a simplistic scheme with two parameters, the length of history window and the frequency of visits, whose values were fixed at default levels. The experiments lasted only a week; however, exploring the effects of changing parameter levels would require longer-term studies.

Combinatorial testing of enhancements to explore synergies and interactions among independent enhancements. In the experiments, we attempted to isolate the features tested to improve the validity of the results and avoid confounding. Enhancements were combined only when they naturally depended on one another; for example, high-information maps were used only with geo-targeted alerts. Controlling for feature presence one at a time was a deliberate design decision. However, when independent enhancements are combined, synergistic effects may amplify the responses, or conversely, feature interactions may dampen or change them. It is also unclear how independent enhancements could be combined *effectively*. More complicated experimental designs would allow testing select features in combination, or, with large samples and longer durations, testing random combinations. Relevant examples include combinations resulting from merging the digest view with other, synergistic high-value features such as precise Geo-targeting, High-Information Maps and Location History.

Better and more comprehensive testing of alert characteristics with more suitable experimental designs. Testing of alert characteristics unfortunately did not yield expected or conclusive results. This was partly due to the inability of the feature-focused experimental design to effectively isolate them, partly due to confounding effects from external factors such as demographics and partly due to the resulting small sample sizes of the factor groups. More detailed and focused studies of alert timing and update alerts are needed to explore their true effects. Studying interactions among alert characteristics themselves, as well as interactions between alert characteristics and enhancement features, is also important since such interactions may turn out to have a non-trivial influence on alert outcomes. These investigations require changes to experimental designs used in this work, however.

Application of compression techniques to higher-precision and smaller geo-targets. In the compression work, we tested our compression techniques with polygon-shaped geo-targets taken from the NWS database and used exclusively for weather alerts. Other applications, especially catering to smaller jurisdictions, may require smaller polygons with higher-precision vertices to support the granularity of geo-targeting needed in certain specialized local contexts. We do not know yet how well our compression techniques work with such polygons and whether we could tune them to improve their performance for specialized applications.

Limited use of URLs in alerts. Some AOs suggested limited use of URLs pointing to short web pages stored and replicated at resources close to the edge of the network, for example at computationally capable base stations serving the alert region. We could explore this idea using CMU's intelligent small-cell cellular network CROSSMobile. Additional details on future research are presented above as part of the Service-level Implications for Rich Media Integration.

Study of human factors in enhancements that rely on specialized user interfaces. Alerts delivered with maps and TTS notifications and the Situation Digest feature depended on new user interfaces not used in the standard WEA service. These user interfaces featured specific layouts and widgets combining sound, maps, icons/symbols and text in particular ways. They were plausible instances of possible implementations and in general represented the simplest solutions imaginable as a first approximation to more complete and better thought out solutions. The user interface undoubtedly affects user experience, which may in turn influence the alert outcomes measured. Therefore, our results were conservatively predicated on human-factors-related decisions made. Future studies should explore the bearing of such decisions in pointed ways to answer to what extent layout, style, icons/symbols, sound, images, maps, text and other elements matter in alert notifications and alert presentation, how to balance their use in alternative visualizations, and to what extent they should be prominent to optimize alert outcomes under different situations.

Inclusion of images in alerts. Possible inclusion of various forms of rich-media content in alerts was a major finding of the AORS phase of our project. However, we only studied responses to maps, external links (URLs and social media tags) and various forms of audio notifications. Some AOs suggested the use of photographs to augment WEA messages, especially in the context of AMBER alerts. The value of adding pictures of persons, as well as locations, properties and other artifacts affected by an emergency, is yet to be studied.

Role of alert creation tools. Our testbed implemented a web-based user interface (the Control Center) to create, schedule and monitor alerts. Our study did not explore issues pertaining to alert creation, however. In particular, addressing situational awareness requires good tools at the source of the alert pathway to allow AOs to easily synthesize CAP-compliant alert messages and manage complex emergency scenarios involving multiple interacting incidents. It would be worthwhile to study the nature and usage of such tools because they potentially affect WEA adoption by AOs.

Situational awareness as a new alerting paradigm. We addressed situational awareness — the ability of an alert recipient to make sense of the information contained in the alert and to correctly infer suitable follow-up actions — using a structured and holistic approach that is demonstrably valuable in complex emergency situations. The real-world implementation of this approach relies on efficient, CAP-based encoding of metadata at the source, new transport mechanisms in the middle and new presentation schemes on the receiving end. We intend to develop and study this idea further, and believe such an integrated treatment could be a precursor to standards recommendations.

Delivery of WEA messages to mobile devices other than smart phones. Having WEA alerts on wearable devices such as smart watches and activity monitors may extend the WEA service's reach. This is an unexplored area expected to have increasing relevance. Once again, the manner in which the information is summarized, digested and/or represented becomes an integral aspect of the potential exploration. Creating mechanisms in WEA to support a variety of carried and wearable devices, as well as situated displays, will become increasingly important as the Internet-of-Things emerges and the value of integrating alerts becomes more apparent.

Re-engineering of WEA for LTE and 5G networks. WEA was engineered under assumptions of network capacities and capabilities reflecting 2G principles — especially bandwidth limitations, centralized network architecture and highly limited capabilities of base stations and their controllers. Twenty years have passed and technologies have advanced substantially. Just as the National Television Systems Committee was created as a television standard in 1941 and eventually was overshadowed by new technologies and new expectations leading to the Advanced Television Systems Committee, so WEA is due for an overhaul. We have a view on both how a modern alerting system could be engineered and ways to make the transition to it incremental. We identify this as an area ripe for additional study, prototyping and experimentation.

Performance Measurements. Currently, there are no effective and systematic mechanisms in place to measure and characterize the end-to-end delays in WEA nor to diagnose system-of-systems problems related to delayed delivery. During our research, we gathered anecdotal evidence of cases where two people in the same location received messages 15 minutes apart. The gap was attributed to differences in phone models or carriers, but there was no means to actually diagnose the problem and rectify it. One AO emphasized that the carriers in general do not report back on the effectiveness of WEA because of liability issues. An end-to-end feedback mechanism such as we have discussed would be one possible approach for tracing delay through the entire service.

Celebrate success stories. As part of educating AOs and the public, WEA's adoption could be bolstered by showcasing stories of how WEA was successful. Anecdotal evidence for AMBER Alerts broadcast through WEA suggests that these have directly helped in resolving at least 12 cases of missing children. Providing some means for collecting and disseminating such stories will raise awareness of WEA and may lead to broader perceived value — helping to minimize opt-out.

Training wheels and growing pains. Ramp-up protocols and best practice guidelines for AOs do not exist. Several of the participants reported being burned in their first attempts to send WEA messages because they did not anticipate the flood of responses. In one case, a department's main phone was completely inundated for a better part of the day.

7. REFERENCES

3GPP. 1999. "3rd Generation Partnership Project; Technical Specification Group Terminals; Technical Realization of Cell Broadcast Service (CBS)." Standard 3G TS 23.041 V3.0.0 (1999-10). http://www.3gpp.org.

———. 2014. "Multimedia Broadcast/Multicast Service (MBMS); Stage 1." Technical Specification Vol. 3GPP TS 22.146. Valbonne, France: 3rd Generation Partnership Project. http://www.3gpp.org.

———. 2015. "Multimedia Broadcast/Multicast Service (MBMS); Architecture and Functional Description (Release 130." Technical Specification 23.246 V13.0.0. Valbonne, France: 3rd Generation Partnership Project. http://www.3gpp.org.

Abowd, G., A.K. Dey, P.J. Brown, N. Davies, M. Smith, and P. Steggles. 1999. "Towards a Better Understanding of Context and Context-Awareness." In *Proceedings of the 1st International Symposium on Handheld and Ubiquitous Computing*, 304–7. HUC '99. London, UK, UK: Springer-Verlag. http://dl.acm.org/citation.cfm?id=647985.743843.

Aloudat, A., K. Michael, and J. Yan. 2007. "Location-Based Services in Emergency Management-from Government to Citizens: Global Case Studies." In *Recent Advances in Security Technology (1 Ed)*. http://works.bepress.com/kmichael/14.

Bourque, L.B., R. Creel, M. Kelley, M. Wood, M. Kano, and D. Mileti. 2013. "An Examination of the Effect of Perceived Risk on Preparedness Behavior." *Environment and Behavior* 45 (5): 615–49. doi:10.1177/0013916512437596.

Braun, V., and V. Clarke. 2006. "Using Thematic Analysis in Psychology." *Qualitative Research in Psychology* 3 (2): 77–101. doi:10.1191/1478088706qp063oa.

Casteel, M.A., and J.R. Downing. 2013. "How Individuals Process NWS Weather Warning Messages on Their Cell Phones." *Weather, Climate, and Society* 5 (3): 254–65. doi:10.1175/WCAS-D-12-00031.1.

CAT. 2015. *Coding-Analysis Toolkit*. University of Pittsburgh. Accessed July 1. http://cat.ucsur.pitt.edu/.

Cohen, J. 1992. "A Power Primer." *Psychological Bulletin* 112 (1): 155. doi:10.1037/0033-2909.112.1.155.

Daly, B.K., M. Gerber, W. Anderson, K. Bhatia, C. Cox, J.S. Davis, T. Dunn, et al. 2014. "Geographic Targeting, Message Content and Character Limitation Subgroup Report." *The Communications Security, Reliability and Interoperability Council Working Group 2 Report*. http://www.rand.org/pubs/external_publications/EP50613.html.

DHS. 2013. "Wireless Emergency Alerts System Enhancement Recommendations." U.S. Department of Homeland Security Science and Technology Directorate.

EC Network. 2015. *CodeRED Mobile Alert*. ECN, LLC. Accessed July 1. https://play.google.com/store/apps/details?id=com.ecnetwork.crma.

Elecont Software. 2015. *eRadar HD - NOAA Radar*. Elecont Software. Accessed July 1. https://play.google.com/store/apps/details?id=com.Elecont.USARadar.

Erdogmus, H., M. Griss, B. Iannucci, S. Kumar, F.J. Diogo, A. Jauhri, and M. Kovalev. 2015. "Opportunities, Options, and Enhancements for the Wireless Emergency Alerting Service." Technical Report CMU-SV-15-001. Carnegie Mellon University.

Erdogmus, H., M. Griss, and Bob Iannucci. 2014. "Enhanced Geo-Targeting Goals and Options for Wireless Emergency Alerts: Alert Originator Requirements Study." Interim Project Report. Carnegie Mellon University.

Gerber, M. Private Communication. 2014. “National Weather Service - Corpus of Polygons Used in WEA Alerts in 2012-2014.”

Golomb, S. 1996. “Run-Length Encodings (Corresp.).” *IEEE Transactions on Information Theory* 12 (3): 399–401. doi:10.1109/TIT.1966.1053907.

Government of Alberta. 2015. “Alberta Emergency Alert.” Accessed July 1. https://play.google.com/store/apps/details?id=ca.ab.gov.aea.

Hagemeier-Klose, Maria, and K. Wagner. 2009. “Evaluation of Flood Hazard Maps in Print and Web Mapping Services as Information Tools in Flood Risk Communication.” *Natural Hazards and Earth System Sciences* 9 (2): 563–74.

Hamilton, B., B. Liu, D. Mileti, J. Sutton, S. Madden, and M. Wood. 2014. “Comprehensive Testing of Imminent Threat Public Messages for Mobile Devices.” 8400 Baltimore Ave., Suite 250 College Park, MD 20742: National Consortium for the Study of Terrorism and Responses to Terrorism (START). http://www.start.umd.edu/research-projects/comprehensive-testing-imminent-threat-public-messages-mobile-devices-mdp.

Hughes, A.L., and L. Palen. 2009. “Twitter Adoption and Use in Mass Convergence and Emergency Events.” *International Journal of Emergency Management* 6 (3-4): 248–60.

Iannucci, B., J.D. Falcao, H. Erdogmus, M. Griss, and S. Kumar. 2015. “From Alerting to Awareness.” In preparation.

Jauhri, A., M. Griss, and H. Erdogmus. 2015. “Small Polygon Compression for Integer Coordinates.” In *Third Conference on Weather Warnings and Communication*. Raleigh, North Carolina.

Junglas, I.A., and R.T. Watson. 2008. “Location-Based Services.” *Communications of the ACM* 51 (3): 65–69. doi:10.1145/1325555.1325568.

Kelley, M.M., M. Wood, M. Kano, and L.B. Bourque. 2010. “Predictors of Household Disaster Preparedness in California: Findings from a Statewide Telephone Survey.” In *Americal Public Health Association 138th Annual Meeting and Expo*. Denver, CO. http://www.colorado.edu/hazards/awards/paper-competition/kelley.pdf.

McGregor, J.D., J.P. Elm, E.T. Stark, J. Lavan, R. Creel, C. Alberts, C. Woody, R. Ellison, and T.M. Keim. 2014. “Best Practices in Wireless Emergency Alerts.” Special Report CMU/SEI-2013-SR-015. Software Engineering Institute.

Microsoft. 2015. “Point Compression Algorithm.” Accessed September 1. https://msdn.microsoft.com/en-us/library/JJ158958.aspx.

Mileti, D.S., and J.D. Darlington. 1997. “The Role of Searching in Shaping Reactions to Earthquake Risk Information.” *Social Problems* 44 (1): 89–103.

Mileti, D.S., and C. Fitzpatrick. 1992. “The Causal Sequence of Risk Communication in the Parkfield Earthquake Prediction Experiment.” *Risk Analysis* 12 (3): 393–400.

Mitchell, H., J. Johnson, and S. LaForce. 2010. “The Human Side of Regulation: Emergency Alerts.” In *Proceedings of the 8th International Conference on Advances in Mobile Computing and Multimedia*, 180–87. ACM. doi:10.1145/1971519.1971551.

Nagele, E. Danielle, and J.E. Trainor. 2012. “Geographic Specificity, Tornadoes, and Protective Action.” *Weather, Climate, and Society* 4: 145–55.

NRC. 2011. “Public Response to Alerts and Warnings on Mobile Devices: Summary of a Workshop on Current Knowledge and Research Gaps.” Committee on Public Response to Alerts and Warnings on Mobile Devices: Current Knowledge and Research Gaps, Computer Science and Telecommunications Board, National Research Council.

———. 2013. "Geo-Targeted Alerts and Warnings: Report of a Workshop on Current Knowledge and Research Gaps." Committee on Geo-targeted Disaster Alerts and Warnings: A Workshop on Current Knowledge and Research Gaps, Computer Science and Telecommunications Board, National Research Council.

OASIS. 2010. "Common Alerting Protocol, v. 1. 2." OASIS Standard. OASIS Emergency Management TC. http://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.pdf.

Oswal, A.K., and J.R. Iyer. 2010. Caching Content at a Cell Site. US Patent 7,801,530, filed 2010, and issued 2010.

SEI. 2014. "Study of Integration Considerations for Wireless Emergency Alerts." Special Report CMU/SEI-2013-SR-016. Software Engineering Institute.

Siegel, A.F. 1994. Practical Business Statistics. Irwin.

Stoddard, R.W. II, J.P. Elm, S. McCurley, T. Sheard, and T.M. Keim. 2013. "Wireless Emergency Alerts: Trust Model Technical Report." Special Report CMU/SEI-2013-SR-021. Software Engineering Institute.

Strauss, A., and J.M. Corbin. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications. https://books.google.ca/books?id=wTwYUnHYsmMC.

Woody, C., and R. Ellison. 2013. "Maximizing Trust in the Wireless Emergency Alerts." Special Report CMU/SEI-2013-SR-027. Software Engineering Institute.

Zaiontz, C. 2015. *Real Statistics Using Excel*. Accessed August 10. Online Book: http://www.real-statistics.com/.

Zickuhr, K. 2013. "Location-Based Services." Pew Research Center. http://pewinternet.org/Reports/2013/Location.aspx.

Ziv, J., and A. Lempel. 1977. "A Universal Algorithm for Sequential Data Compression." *IEEE Transactions on Information Theory* 23 (3): 337–43.

Appendix A1. Alert Originator Requirements Study (AORS) Interviews — Detailed Study Design and Findings

A1.1. Goal

The goal of the AORS in *Goal-Question-Metric* format (Basili, Rombach and Caldiera1994) is expressed as follows:

Validate	previous findings and recommendations regarding the viability, limitations, advantages and use of the Wireless Emergency Alerts (WEA) service (G1)
and	
identify	impediments and opportunities regarding the adoption of the WEA service by alert originators (AOs) (G2) and acceptance of the WEA service by the public (G3)
for the purpose of	prioritizing, developing and evaluating enhancement goals and options for future wireless broadcast services (G4)
with respect to	increased geo-targeting specificity (G5) of alert messages and improved relevance (G6) to their recipients
from the perspective of	AOs (G7)
in the context of	emergency alert systems and services used by different jurisdictions, agencies, and response communities throughout the nation (G8).

A1.2. Assumptions

A1 AOs have a sense of the needs of public regarding the consumption of emergency alerts and how the public feels about them.
A2 AOs have a sense of the weaknesses and strengths of the WEA service.
A3 AOs have a sense of what it would take to address the weaknesses of the WEA service.
A4 AOs have a sense of how to leverage the strengths of the WEA service.
A5 AOs use (or are willing to use) WEA messages only in limited circumstances due to current limitations of the WEA service and the delivery mechanisms underlying them.

A1.3. Research Questions

RQ1	Is more precise/granular geo-targeting likely to improve the effectiveness of WEA messages? (G1, G5)
RQ2	Is increasing the relevance of WEA messages to recipients based on the recipients' context likely to improve the effectiveness of the WEA service? (G1, G3, G6, role of context)
RQ3	Is better control by the recipient of which messages they will receive and which messages they wish to disregard likely to improve the effectiveness of the WEA service? (G1, G3, G6, interest targeting)
RQ4	Is having a feedback mechanism (implicit or explicit) in WEA delivery from the recipient back to the AOs likely to improve the adoption of the WEA service? (G2, G7)
RQ5	Is the use of better alert creation tools that help create meaningful and targeted alert messages likely to improve the adoption of the WEA service? (G2, G7, multimedia support)
RQ6	Is relaxing the length limitation of a WEA message likely to improve the effectiveness of the WEA service? (G1, G3, message length)
RQ7	Do different kinds of AOs have different needs and reasons for adopting and using WEA messages? (G2, G7, G8)
RQ8	Are certain improvements to WEA more important than others in terms of their potential to improve the effectiveness of the WEA service? (G2, G3, G4)

A1.4. Participant Selection

The data collection portion of the work involved in-depth interviews with a set of subjects selected from the sample population of about 600 AOs assembled from a combination of sources. A substantial effort was made to balance the subject selection and maintain diversity of location and scale. We included AOs from entities that operate on local (e.g., city, county), state and national levels, and sought representatives from all regions of the country. We included participants dealing with a range of emergencies, including weather related ones (e.g., hurricanes, tornadoes, flooding), earthquakes, fires, shootings, as well as cases of abducted children.

Initially we set out to include as many AOs that had adopted WEA as possible. This turned out to be a very difficult task, since WEA adoption is still in the very early stages; finding subjects with requisite direct experience proved elusive. Therefore, we chose to also include AOs who have direct experience working with other emergency alert systems, such as AlertSCC used in Santa Clara County, California, but are open to using the WEA service in the future. A number of local emergency services organizations to which some of these AOs belong are already set up to issue WEA messages, but are currently either rarely using the WEA service or not using the WEA service at all. Quite a few are in the process of becoming authorized AOs, or are prepared, but have not had the opportunity or need to issue WEA messages. We augmented these with informal, unstructured interviews with five other local AOs and experts who were willing to talk to us.

A1.5. Profile of Participants

The participants included AOs who were interviewed using the unstructured approach (the last five entries in italics). All participants held senior roles in various emergency management organizations with mandates over different jurisdictions (national, state, county, local). The states represented in this sample were California, Kansas, Massachusetts, Texas and Florida.

The participants' titles spanned Senior Coordinator of the Office of Emergency Services (OES), Director of OES, Program Lead of Emergency Communications, Emergency Management Liaison, Assistant Director of Emergency Management, Preparedness Coordinator, Director of National Center, OES Coordinator, Emergency Management Specialist, Emergency Operations Center (EOC) Coordinator and Head of Protective Services. Their affiliations spanned Santa Clara County, California, City of Palo Alto, California, National Weather Service, Harris County, Texas, Johnson County, Kansas, State of Florida, State of Massachusetts, National Center for Missing and Exploited Children, the National Aeronautics and Space Administration (NASA) Ames Research Park, and Department of Homeland Security/Federal Emergency Management Agency (FEMA).

A1.6. Semi-Structured Interview Protocol

For the core eight subjects, the interviewers used a semi-structured approach, combining questions and observations. This field study approach is called *Contextual Inquiry* (Ross 2014). This allowed us to understand the context in which AOs work with emergency situations and alerts, as well as uncover any related needs and pain points. Two of the interviews were conducted *in situ* where we were able to observe AOs at their desks while they illustrated how they dealt with an emergency situation or created emergency alerts. The remaining semi-structured interviews were conducted remotely through Skype or over the phone.

During the semi-structured interviews, a mixed interview protocol was adopted to elicit both targeted and emergent information. Targeted questions were aimed at validating or refuting our research hypotheses and solution ideas. A script composed of a sequence of primary questions and follow-up questions guided interviews. Diverging from the script was allowed to capture emergent information by letting the subjects drive the interviews. For example, one of the insights uncovered due to this kind of flexibility revealed the importance of connecting social media to WEA.

A1.7. Open Interview Protocol

The five additional interviews followed an open-ended format; they did not use a pre-determined structure with a set of questions and did not take advantage of the Contextual Inquiry approach. Four of these interviews were *in-situ* and conducted in the San Francisco Bay Area. One was remote and conducted over the phone. Two open interviews were conducted together. In each case, either the participants specifically requested an informal meeting with a broader scope or we had not yet formulated the interview protocol, but accepted to meet to not miss the opportunity. During these meetings, the research team also discussed collaboration opportunities and additional matters with the participants, extending their scope and giving them a multiple purpose. The relevant observations were nevertheless recorded and summarized using the same approach as the semi-structured interviews.

A1.8. Data Analysis Methodology

The data from all interviews were analyzed using the Grounded Theory approach (Strauss and Corbin 1998). Grounded Theory is a qualitative research methodology that operates effectively in a reverse fashion from traditional hypothesis-based scientific research. Unlike traditional models of research that begin with a hypothesis, Grounded Theory is an exploratory method that is centered on data. The theory is constructed through the analysis of data, thereby making the findings and theories developed within a study “grounded in the data.”

The key goal of the Grounded Theory study is to discover a research participants’ main concern and how they continually try to resolve it. Researchers must repeatedly ask themselves, “What are the main problems the participants are grappling with, and how are they trying to solve them?” In most behavioral research endeavors, persons or patients are units of analysis, whereas in Grounded Theory, the unit of analysis is an *incident* – such as a statement made by a participant during an interview. Throughout the interviewing process, significant statements made by research participants are recorded through a process known as “noting.”

The empirical data collected through these notes are then sorted and clustered to identify patterns. From these patterns, researchers generate themes that explain the common ways in which the research participants in the study resolve their central concerns. These concerns are then related to research questions. The results of Grounded Theory are not a set of statistically significant probabilities, but a set of statements about the relationships among concepts, or an integrated set of conceptual hypotheses developed from empirical data. Because deviations from the scripted questions are allowed to capture emergent information, often not all questions from the initial set are addressed during an interview, and hence quantitative results on the questions are typically not gathered or reported. Such quantitative results are best obtained through other research methods, for example, through a follow-up survey.

A1.9. Data Collection

In the interview protocol, a debriefing session followed each interview. The researchers who participated in an AO interview met as a group immediately after the interview session to capture what was learned. During these discussions we captured key themes, significant quotes and data points from our conversation with the AO. We captured this data on Post-It® notes. The tactile quality of Post-It® notes made it easy for the team to scan pieces of data and engage with them in a collaborative fashion. Noting also provided a shared sense of ownership in the process, making it easy to build upon the ideas of others. After each interview, one researcher was assigned to review the audiotapes from the AO discussion to capture any additional relevant data points and generate an interview summary for each participant.

A1.10. Data Aggregation

Next, after all the interviews were complete, the entire research team met during a half-day data synthesis session to review the data that had been collected on the Post-It® notes from the interviews and generate additional notes by going through the topline summaries. Observations from the eight semi-structured and five open interviews were aggregated at this point. The team gathered the notes and clustered the data into themes.

Common quotes and ideas were clustered together. The tangible nature of this process allowed the team to synthesize the data collaboratively, enabling a shared understanding of the key themes. The process enabled the team to visualize and sort the common statements and ideas that were communicated by the AOs interviewed.

After several hours of discussion and debate, the team arrived at five *Primary Insights* that we collectively believed were emergent and important and that directly bear onto the research questions. All additional insights were captured in the *Secondary Insights* list.

A1.11. Primary Insights

The AORS interview results are grouped into five Primary Insights and a number of Secondary Insights. The Primary Insights cover topics and themes that were most strongly articulated by the AOs, and they have a more direct bearing on the research questions. We discuss each of the Primary Insights through inclusion of specific quotes from the participants and examples from the emergency messages sent to the public. Both the Primary and the Secondary Insights are later used in the discussion and recommendations to derive implications for future studies, in particular for the AO Survey and Trial components of the Carnegie Mellon University (CMU) WEA Project.

Insight 1

I1	90 characters are not enough to convey meaningful information to the public.

The most common problem reported by the participants is working within the current 90 characters constraint of the WEA service. All participants interviewed stated this stringent character limit was simply not enough to convey relevant messages to the general population. The AOs emphasized the need to include information in the alert messages pertaining to the originator, relevance, expected action and external sources, which the 90-character limit hardly allow.

"We can't just issue a casual message. To get people's attention, probably one of the most limiting factors of WEA is the 90-character length. So you really can't have a full-blown explanation of what the emergency is in 90 characters. Now that the system is capable of more… it is an educational process to [show] the public what [WEA message] is. It is not a text message, it is not an email… it arrives on the carrier's frequency on the phone… With this biggest issue we have, we work with local authorities to just get the people's attention." (BKN)

"You throw in the time, type of warning, the action that people need to take, to check media… that is all that you can put in there..." (MGR)

"[In relation to missing children with autism]… because children with autism are attracted to larger bodies of water… This would also need more than 90 characters to explain it." (BHR)

Some of the participants worked with other systems where the message length was less constraining, most notably the full "America's Missing: Broadcast Emergency" (AMBER) Alerts system with the 160 character limit and Twitter with the 140 character limit. Their experience suggests that even those systems can be too limiting to include important details:

"The next challenge is the 90 characters and that is something everybody is aware of. I have brought it up with the wireless industry… we have been working with wireless industry since 2004 for AMBER Alerts when people had to sign up for that and we have lots of historical information… It used to be 160 characters. And we could barely get enough information into that." (BHR)

Because there is little awareness by the public around wireless emergency alerts, WEA messages carry the additional burden of both educating users as well as conveying information about the emergency at hand. Most participants in our study concurred that they believe a key reason users opt out of the service is because they are simply confused when they receive a WEA message for the first time and are unsure how they are supposed to follow up. Extra information in the message may address this problem, which the current message length limit does not permit.

In addition to increasing the amount of content expressible solely by text, the most frequent observation by the participants was the need to include pointers to supplemental information or photos. This need is illustrated by an example of a Twitter message, which includes both the Uniform Resource Locator (URL) for the follow up and an image. Such presence of mixed media in alerts is related to both message content limitation and the intent of the alerts.

Closely related to message length is message structure. During the interviews, one AO stated that there are three components to a good alert message:

"Key elements that you need in an effective message are: clarity from whom it is coming… that by itself is challenging (so people know that it is not spam)... Next is: what is going on, and why I should care?… what is in it for me?… the last one is: what is the expected action, what do I do with that?… There are cases when you have to [send] a mini URL" (KDR)

The length limitation makes it difficult to properly articulate each of these components.

Insight 2

I2 Geo-targeting of WEA messages is seen as the long sought out goal by AOs.

Most of the experts interviewed believe that geo-location functionality provided by modern smartphones will enable more precise control over the dissemination of WEA messages. This functionality will allow a recipient's smartphone to suppress the alert message if the phone determines that the recipient is outside the targeted geographical area. For such on-device filtering to be possible, the WEA message must carry the geo-target with it. While the current approach of sending WEA messages to users in a targeted area with cell phone coverage is seen as a vast improvement over most landline emergency systems, many participants underscored some of the drawbacks of the pure textual cell broadcast approach.

"The reason why [we adopted an app for emergencies] is because you can draw a circle or a polygon on a map and you can send alerts to that area based on [where] people … are. So people are not getting a coastal flood warning if they live in the western part of the state." (CBE)

"How do you alert the public without over-alerting them… with WEA technology that uses cell broadcast… Because it is a broadcast, the radio signal can bleed over [to untargeted areas]. So you can be outside of the threat area and receive the warning… so it is not relevant to you… You systematically increase the likelihood of desensitizing people ... This phenomenon is most common in the rural locations…" (MGR)

"… We have our ability to send WEA [messages], we have processes in place, but we have not used it because of the limited range of capabilities of that technology… Our challenge is that our county is [large] with [many] cities… That is a lot of geographic space and a lot of population…" (FSZ)

Utilizing the geo-location functionality embedded within most smartphones sold today would allow the WEA service to more precisely target people within a designated area. Most participants interviewed believe that utilizing geo-targeting based on smart-phone geo-location capabilities will rectify many of the key challenges faced by the platform.

The special situations include AMBER Alerts for which a larger area such as a whole state still needs to be alerted for cases of abducted children. Nevertheless, issuing AMBER Alerts for cases of certain missing children, such as autistic children wandering off, would also require only coverage of a limited area.

"In Maryland it is going to hit the whole state. In New York they have broken it down to the areas… In New York there are a lot of abductions that are on foot, then we could use it on the county level… But if we could use it for missing child with autism that wander off (last year we had 14 who died)… then we would do WEA just for a neighborhood or a slightly larger area… [or around large bodies of water] because children with autism are attracted to larger bodies of water." (BHR)

The most effective ways of describing the targeted area is a key question. The participants expressed their desire to define these targeted areas in a variety of ways, such as based on regions on a map, distances from the source of the emergency (as a circle) or through the provision of an arbitrary polygon that defines the emergency area:

"… We need to leverage the capabilities of those smart phones. For example… with these vertices or polygons (that describe weather threats), if we could push [WEA messages] [to the phones]… some studies show that this could improve the public response to them." (MGR)

Geo-targeting represents only one dimension of context-aware filtering of information. The reports from participants suggest that there may be other important factors, such as time of day or other elements that are specific to the recipients' context. For example, we learned from one AO of an incident where many recipients of the WEA messages complained about having been disturbed with a message sent in the middle of the night. Nevertheless, geo-target-based filtering emerged as the top concern pertaining to context-aware filtering on the device.

Insight 3

I3	The WEA service needs to interface with social media to be relevant.

Given the rapid adoption of social media services such as Twitter, Facebook and Nextdoor (a neighborhood-based social network that is rapidly gaining popularity), participants expressed great interest in connecting emergency alerts and the WEA service to relevant social media. It was illuminating to hear the extent with which participants interviewed already used such social media channels within the context of their emergency management role. The rise of social media has certainly altered the public's expectations about the timeliness of emergency alerts:

"We have a wide range of use of public alerts and notifications from natural disasters to hijacked cars…to VIP visits… It is not just earthquakes or crimes, it is all hazards and the fact that we are trying to cope with the proliferation of the social media and the change in the expectations of the public… you expect to be notified as soon as possible. Ten years ago your expectation was, if something bad happened an hour ago, it would be nice if something came over the land-line. But now you expect something immediately." (KDR)

"We appreciate the digital environment. We appreciate social media. We have a mobile-friendly web site, but Facebook and Twitter is one way people can access emergency alert information on their mobile device." (FSZ)

The participants saw social media outlets as valuable communication channels and easy, lightweight mechanisms for conveying information to the public. Social media are also being used in non-emergency situations, for example to engage the public and increase overall preparedness and awareness.

"We are trying to get to people where they live. [That is] the reason we adopted social media, and we have been an early adopter… Our police chief did first-ever "tweet along," our lieutenant <Name>

would drive with the chief, and every time he would make a traffic stop or pull somebody over, there would be a tweet about it. And it seems kind of like a gimmick, but since [then], there has been a bunch of agencies that have done it. We are doing it as a means to engage populations where we feel there is a gap in outreach for public safety, especially youth and parents of young children… typically people 30 or younger or parents of young children because they are very busy…" (KDR)

"Two weeks ago we had a fire on Foothill, and Page Mill Road was closed from 280… At the scene we had a conversation and we decided there was little danger of fire spreading, and so there was no imminent danger, more dealing with traffic ... So we decided to send it out via Twitter. Not imminent threat to life, more of a headache…" (KDR)

"… In the past when something was going on, if we at the state wanted to let the public know…we had two ways to do it: radio and TV (traditional media) or send messages or make calls to the local authorities… those were traditional things… it was difficult to get out the message to the public… social media changes things because it cuts out the middleman in our messaging and what we want to put out… it is very good to get out messages unfiltered, and media picks it up very quickly." (CBE)

"In social media, about 95 percent of our effort is [directed to] Twitter and Facebook…We use the same account for both preparedness and alerts…We see retweets…" (CBE)

One of the AOs also reported on the increased adoption of community-based social networks, such as Nextdoor and speculated that such networks could play a greater role as alternative emergency information dissemination outlets.
AOs imply that some users may want to consult multiple information sources before trusting the validity of news or information. Although the WEA service is a broadcast communication system reserved for alerting the public to emergency situations, it lives in the context of an ecosystem of broadcast and interactive communication systems. There is potentially great value in allowing the platform to hook into existing interactive communication systems. This would build on existing and well-studied user behavior and likely increase the efficacy of the system. However, the concern from wireless carriers is that allowing interaction with other wireless Internet services will negatively impact network load and counter one of the primary goals of using the wireless broadcast technology in the first place – that of reduced impact on the wireless infrastructure.

Several AOs in the study expressed belief that, despite the required investment in resources for building presence and participating in the social media, these communications channels are extremely valuable because they allow the AOs to build relationships and trust with the public over time.

"There are those that have an instinct to back away from social media, especially Twitter because it involves resources, it involves engaging with the public, it involves that two-way interaction. But Twitter is the best thing to happen in emergency management and emergency management information...Twitter for us is an extraordinary tool. I don't have to write seven or eight press releases during an emergency. I all have to do is create the bullet point in 140 characters. And I would rather do that 20 times a day because it allows me to give the public information in real time." (FSZ)

The AOs speculated that with a trust relationship built over time and a two-way dialog established, when an emergency happens, the public is more likely to follow official instructions. This work of building and maintaining trusted relationships never stops and it also includes sending out reassuring messages, such as in cases of tragic events.

The interviews with AOs identified that social media is currently an integral part of how emergency personnel communicate and interact with the public at large. While social media has some drawbacks, these drawbacks are clearly outweighed by the benefit of the network effect that is gained through these platforms, as well as their ability to effectively and efficiently convey information to the public. It is important that further iterations of the WEA service do not merely replicate what social media platforms are already achieving and attempt to create an alternative platform. Instead, we recommend considering how the WEA service can co-exist within the constellation of social media platforms and how social media's inherent features and functionality can be leveraged within the WEA service.

Recent research supports the AOs' perceptions regarding the role of social media. Liu et al. (2014) investigated the effect of social media on disaster response and concluded that social media increased participants' intended likelihood to seek additional information and take action. These authors stated that "participants were more likely to seek further information from Twitter when the initial disaster information was in the form of a tweet than a web page."

Insight 4

I4 There are two distinct conceptual models of WEA: a mere warning alarm or a richer media application with follow through.

A key unresolved issue that emerged in the AORS was around the core conceptual model of WEA, which impacts how the users of the system envision it evolving. Some participants viewed the system – either because of its limitations or because of its intent – as an alarm bell with minimal functionality, similar to a siren.

"In its current state, WEA is like an alarm bell technology. It alerts you that there's an emergency and to go find out about the emergency from other sources." (BHR)

"Now, with 90 characters, all WEA can realistically be is a bell-ringer technology." (FSZ)

"There are a lot of legacy devices out there. For example, there are people in my family who have a flip phone. If you have WEAs with links, images or attachments, not all devices are provisioned to access that kind of information." (FSZ)

Others see the current WEA service as the early beginnings of a rich interactive communications platform that will eventually evolve to include maps, images and links to additional information that is curated by emergency management programs. Enhancements in that context could also entail integrated follow-up, feedback and closure mechanisms that generate a two-way trail for better traceability and response measurement.

"I've been begging to just get a link (in the WEA message) to www.amberalert.gov and we would maintain the data on that site… The wireless industry is concerned that everybody would go to that and bring down the system. This is the debate we have right now and the argument we use is that people already go to amberalert.gov anyway since they don't have enough information (in the WEA message)." (BHR)

"I think in the future with WEA, you'll definitely see more text. I'd really like to see some sort of graphics. I'm pushing really hard for graphics and being able to push the vertices to the device through a graphic. Now we just do county codes... I would also like to see more ability for the user to configure WEA [messages]. For example, I'd really like to see engagement from the disability community." (MGR)

"One of the challenges we have here in [this state] is that on any given day, we have thousands and thousands of tourists in our state. Now most of these tourists don't know what county they are in, or the highway numbers. So characters alone don't really work for these tourists." (BKN)

In particular, the rich media messages could provide for extending the purpose of AMBER Alerts to missing persons of all ages.

A clear challenge of the WEA service is that it is currently designed to be a broadcast only (one-way) communication system that is delivered to a device that is conceptually understood to support two-way communication. Further, the WEA service is situated in a technology landscape in which users expect interaction. While the evolution of technology is a moving target, this insight points to a need to determine a perspective and strategy regarding the *mental model of WEAs*. Since AOs will be an essential element in the evolution and growth of the WEA service, it is important that they have an understanding of how FEMA and other stakeholders view the system and its evolution, and then develop a collective vision.

Another aspect of interactivity is follow though. For WEA alerts to be beyond an alarm bell, receivers of WEA messages may need to know how an emergency situation evolves and when the emergency is over. At present time, WEA does not provide for linked "clear all" type of closure messages when the emergency is over.

Insight 5

I5 Better outreach and education for both the public and potential AOs will improve acceptance and adoption of the WEA service.

One of the key challenges our research team faced throughout this work was recruiting proficient and active WEA users. While it was easy to identify and recruit participants with positions in emergency management, most did not use WEA and were uncertain as to how it worked. Additionally there was concern that if they did use the system, the public response to errors that are an inevitable part of learning a new technology had the potential to create a public relations firestorm for the organizations that they are a part of.

"I recall when WEA was initially rolled out and I was kind of excited about it. And then it rolled out with no outreach… it was just…there. I don't think there has been enough push to make both the receivers and the providers aware of what WEA is and what it can do. You've got to talk this thing up, otherwise you risk people just turning it off and people in emergency management not using it." (BK)

"The beginning of WEAs and AMBER Alerts was very painful for me. We were under the gun to get it out there, and a lot of people complained about this obnoxious tone that was created by the National Center. It's the same tone that goes over the television and radio – but people were not expecting that tone to come out of their phone. There could have been a much better public awareness around what the tone meant. When we first launched WEA, I had to do 70 press interviews in one week because people didn't know what it meant. Now people complain about not getting an alert." (BHR)

Many of our participants believed that there exists a substantial need for outreach and education for the general public. They explained that the benefits of the WEA service, as well as the appropriate action that should be taken upon receiving alerts, remain elusive to the majority of the general population.

"Outreach is a huge part of this. A couple of years ago, I was in Las Vegas and I was the host of a panel discussion and I said we've got enough engineers in the room. What we don't have enough of are people to do outreach… We can cover what we need to do with technology. It can be done. But unless we reach out to the general public, there's going to be a lot of confusion. I was banging my fist on the table about this point. Did they listen? Ah… I don't think so. I think more could be done." (MGR)

"AMBER [Alerts]… are very, very different… When you send WEA alert with that loud tone to get people's attention… it is designed to create alarm in people to let them know they are in danger and get them to take immediate action because they are in danger. The AMBER Alert is totally different. The people you are alerting are not in any danger, but we believe they may have information that can save the life of a child. That child is in danger. So that is why there is a major difference… People don't even understand what this tone is and why it is coming from the phone... There could've been much better public awareness … But some awareness is there now." (BHR)

Public education may also address fears of privacy invasion. Many members of the public expressed initial concerns regarding their invasion of privacy based on the erroneous assumption that geo-targeting meant that the WEA service was tracking their location. If geo-targeting is performed through filtering on the smartphone and the public is educated about this capability, citizens' fears about the WEA service monitoring phone locations may be alleviated.

While education outreach may fall outside the scope of the CMU WEA Project, it remains an important insight gained from this study. A consistent message that we heard from the participating AOs is the "build it and they will come" marketing approach for the WEA service will likely prove unsuccessful. There is a wealth of technical knowledge and development that can be brought to bear in future iterations of the WEA service. It is our belief that marketing WEA and educating both the public and AOs about the benefits and functionality of WEA will be as, if not more, essential to the success of the WEA service.

A1.12. Secondary Insights

The Secondary Insights listed here cover an additional number of topics and issues that were articulated during the interviews. As noted before, the Secondary Insights were those that were not as prominently emphasized by the AOs interviewed as the Primary Insights. Thus the research team did not deem these as significant as the Primary Insights, but considered them still worthy of mention for completeness.

- **Language issues.** Two AOs in our study reported the need to address multi-language and translation issues for any type of an emergency message that is sent out. Tackling multiple languages may be a good application for context-aware filtering. If alerts could be issued in multiple languages, the phone could be set to choose the one in the user's appropriate language.

- **Technical issues**. While the public is receptive to WEA, people do not want to be needlessly interrupted. The technical problems in the system implementation cause repeated messages for some recipients. One participant commented on the general attitude towards AMBER Alerts and cases of complaints related to repeated messages due to technical problems with the system:

"99.9 percent of the people out there want to be involved and want to be that person who has the information, but they don't want to be annoyed every five minutes." (BHR)

- **Performance measurements**. Currently there is no way to measure the technical performance of the system. We heard anecdotes of cases where two people in the same location received messages 15 minutes apart because their phone models or carriers were different. One AO (WWR) emphasized that the carriers in general do not report back on the effectiveness of WEA because of liability issues.

- **How effective is WEA?** The overall effectiveness of the WEA is also very difficult to measure. Our participants mostly heard from the vocal minority that complained about the problems with the WEA messages.

- **Celebrate success stories.** WEA's adoption could be bolstered by more "success stories." Anecdotal evidence for AMBER Alerts broadcast through WEA suggests that these have directly helped in resolving at least 12 cases of missing children. However, no systematic evidence is being collected.

  "Some of the most successful stories that we get with WEA is with tornadoes and when people are in the location where they would not normally receive an alert, like in a church [that is in immediate danger]…" (MGR)

- **Training wheels and growing pains.** Ramp-up protocols and best practice guidelines for AOs do not exist. Several of the participants reported being burned in their first attempts to send WEA messages because they did not anticipate the flood of responses. In one case, a department's main phone was completely inundated for a better part of the day.

- **My phone knows me.** Context awareness of the personal phones and their extensive knowledge of their users behavior and prior history is something that can be better leveraged.

A1.12. Answers to Research Questions

RQ1 Is more precise/granular geo-targeting likely to improve the adoption and effectiveness of the WEA service? *Collected data strongly suggest that geo-targeting will improve the adoption and effectiveness of WEA messages. Supporting evidence is discussed in detail in insight **I2**.*
RQ2 Is increasing the relevance of WEA messages to recipients based on the recipients' context likely to improve the adoption and effectiveness of the WEA service? *The evidence related to this question is inconclusive. This question is discussed in insight **I2** only in relation to geographical context. There is no indication that the AOs believe general context awareness is a central issue, however. Certain elements of context (e.g., timing of messages) were brought up and could be worth investigating further.*
RQ3 Is better control by the recipient of which messages they will receive and which messages they wish to disregard likely to improve the adoption and effectiveness of the WEA service? *We have no evidence either way because currently the recipients only have a mechanism to completely opt-out. This question was not probed deeply with the participants since it is more suitably explored through studies conducted with the public.*
RQ4 Is having a feedback mechanism (implicit or explicit) in WEA delivery from the recipient back to the AOs likely to improve the adoption and effectiveness of the WEA service? *Collected data only weakly supports this hypothesis. The absence of ways to measure the technical performance and the effectiveness of WEA messages is discussed in the Secondary Insights.*

RQ5 Is the use of better alert creation tools that help create meaningful and targeted alert messages likely to improve the adoption and effectiveness of the WEA service? *Collected data supports this hypothesis. Alert creation appears is a pain point for AOs, as discussed partially in insight **I1**.*
RQ6 Is relaxing the length limitation of a WEA message likely to improve the adoption and effectiveness of the WEA service? *Collected data strongly supports this hypothesis. The topic is discussed in detail in insight **I1**.*
RQ7 Do different kinds of AOs have different needs and reasons for adopting and using WEA messages? *Collected data supports this hypothesis to some extent. The evidence is subtle, but present in several insights and it is best articulated for AMBER Alerts (for example, vis-à-vis weather- or disaster-related alerts) in insight **I5**.*
RQ8 Are certain improvements to WEA more important than others in terms of their potential to improve the effectiveness and adoption of the WEA service? *Collected data strongly supports this hypothesis. Based on the insights **I1** and **I2**, message length, inclusion of URLs for follow-ups and geo-targeting are the most important aspects that need to be improved in order to increase WEA adoption for AOs. Also, the public awareness and smart utilization of social media are important for increasing adoption for the general public (insights **I3** and **I4**).*

A1.15. Appendix A References

Basili, V.R., H.D. Rombach, and G. Caldiera. 1994. "Encyclopedia of Software Engineering." In *The Goal Question Metric Approach*. Wiley.

Liu, B., J. Fraustino, and Y. Jin. 2014. "How Disaster Information Form, Source, Type, and Prior Disaster Exposure Affect Public Outcomes: Jumping on the Social Media Bandwagon." *Journal of Applied Communication Research*, December, 44–65.

Ross, J. 2014. "Why Are Contextual Inquiries So Difficult?" *UXmatters*. http://www.uxmatters.com/mt/archives/2012/06/why-are-contextual-inquiries-so-difficult.php.

Strauss, A., and J.M. Corbin. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE Publications. https://books.google.ca/books?id=wTwYUnHYsmMC.

APPENDIX A2. ALERT ORIGINATOR REQUIREMENTS STUDY (AORS) INTERVIEW QUESTIONS

ID	Interview Question Category
0	**Demographics and Ice Breakers**
1.1	Name
1.2	Role
1.3	Organization
1.4	Years of experience as an Alert Originator (AO)
1.5	Tell us about your experience/background as an AO.
1.6	Tell us briefly about your work and walk us through a typical day.
1.7	What do you like the most about your work?
1.8	If you could change one thing, what would that be?
1	**Background on Emergency Alerts**
1.1	Do you currently issue emergency alerts?
1.2	How frequently do you issue alerts?
1.3	What types of alerts do you issue? What alert services do you use to issue them?
1.4	What kind of experience do you have with Wireless Emergency Alerts (WEA)?
1.4a	*Follow-up: What are your general impressions about WEA?*
1.5	Which emergencies alert services/channels are you using to disseminate alert messages?
1.5a	*Follow-up: Why? What are the advantages of using these alert services/channels?*
2	**Creation of Emergency Alerts (if using WEA, focus on WEA message creation)**
2.1	Could you show us [or tell us if not in operation center] us how you react to an emergency situation that might result in an alert, and how you create an emergency alert?
2.2	Which tools are you using to create emergency alert messages?
2.3	*Follow-up: Is there a preferred tool? Which one?*
2.4	Do you think the current alert creation tools you are using are intuitive enough?
2.4a	*Follow-up: What are some of the features in the existing alert creation tools that you really like?*
2.4b	*Follow-up: If not, what are the features you want it to have?*
2.5	Do you need your tools to provide more support in constructing alert messages? Elaborate.
2.5a	*Follow-up: Are they sufficient in their capabilities to target to appropriate geographical area and to appropriate people?*
2.6	Do you need your tools to provide more post-issuance support (tracking, archiving, analysis)? Elaborate.
3	**Appropriateness of Existing Alert Message Formats (e.g., Common Alerting Protocol - CAP)**
3.1	Do you feel that you are able to include all necessary details in an alert?

3.1a Do you feel that you are able to include all necessary details in a WEA message?

3.1b Follow-up: Is there any information that you often feel is not properly captured in a longer alert message, but would be beneficial to be present in a WEA message?

3.1c Follow-up: Is there any information that you often feel could be captured in a longer alert message, but not captured in a WEA message?

3.2 Do you use specialized stored templates when creating alert messages?

3.3 Are you aware that you can either let the Integrated Public Alert and Warning System (IPAWS) create WEA message text from the data you provide, or you can write your own WEA message text?

3.3a Follow-up: Which of these methods (automatic construction of WEA text or constructing your own WEA text) do you use or would you prefer?

3.4 Could alert messages include references to other sources where the public could obtain additional information? If so, what kind of sources?

3.4a Follow-up: Ask the same question about WEAs. If asked, give examples: a Uniform Resource Locator (URL) to issuing Emergency Management Agency (EMA) site; social media, e.g., Twitter hashtag.

4 Adoption of WEA

4.1 Do you believe fewer AOs are using WEA than one would like or expect?

4.1a Follow-up: If so, what are the reasons?

4.2 When do you use/prefer non-WEA alerting services over WEA? (e.g., Blackboard Connect) Why? How do they compare in experience and effectiveness?

4.3 If we could implement a single improvement to the current WEA service, or address one limitation, what should that be?

4.3a Follow-up: Can you convey the critical information needed for an alert within the 90-character message limitations of WEA?

5 Public Feedback Regarding Alert Messages (focus on feedback on WEA if WEA is used)

5.1 Do you keep track of how many recipients receive and read alert messages? [Very few services provide this tracking capability. Reverse 911 systems can track how many calls are answered. Email can track message receipt and opening. No other methods offer tracking.]

5.2 What kind of feedback/statistics do you get on effectiveness and appropriateness of alert messages?

5.2a Follow-up: What kind of feedback/statistics would you like to gather to gauge the effectiveness and appropriateness of alert messages?

5.3 Do you hear about the public wanting to opt out of WEA messaging?

5.3a Follow-up: How pervasive is the desire to opt out? What are the reasons?

5.4 Do you hear about the public wanting to customize the way they receive messages or having more control over the messages they receive? If so, what kind of messages and under what circumstances do they want to receive them?

5.4a Follow-up: Do you know how many recipients want to customize their alert profile or message receipt preferences?

5.5 Do you have information on whether the recipients act on the messages received?

5.5a Follow-up: Is it the majority or minority of the recipients who act on alert messages received? If minority, why don't they?

5.6 What differences do you see in how people react to WEA messages vs. alert messages received through other IPAWS channels?

6 Geo-targeting and Interest Targeting of Alert Messages

6.1 What geographical range do emergency situations handled by your organization tend to cover? What geographical range do the alerts issued by your organization tend to cover? (Street/Block/Neighborhood, Multiple Streets/Blocks/Neighborhoods, Town/City, Multiple Towns/Cities, County, Multiple Counties, Region, State, Multiple States, National)

6.1a Follow-up: Is the geographical coverage the same for WEA messages as it is for alerts delivered through other channels?

6.2 Is geo-targeting precision a problem?

6.2a Follow up: If we made geo-targeting much more precise than is currently possible for recipients with Global Positioning System (GPS) capabilities, would that help?

6.3 Currently, the minimum required geo-targeting for WEA is at the county level. Is this sufficient for your needs?

6.3.a *Follow-up: If this level of geographical specificity is not sufficient, what would be sufficient?*

6.4 Would you initiate more WEA messages if they were better targeted?

6.6 Are you working or have plans to work with wireless carriers regarding more precise geo-targeting? If so, elaborate.

6.7 Currently, a WEA message can only be received in and around the affected area related to that message. Should the public also be notified when a WEA message is issued to their home area or a user-specified area of interest, even if they are physically outside that area at the time the alert message is broadcast?

7 Contextual Relevance of Alert Messages

7.1 Do you think targeting based on recipients' interests, profiles, patterns, preferences and situation, and not just location, will increase effectiveness and adoption of WEA alerts?

7.1a Follow-up: What kind of contextual or situational factors should/could be considered for effective targeting?

8 Additional Factors and Closing

8.1 Are there any issues or opportunities that influence the use of an alert service (either from the public's perspective or from your perspective) that we have not covered?

8.2 Is there anything else that we should know?

8.3 Would you be willing to discuss this topic further with us if we have any follow-up questions?

Appendix A3. Alert Originator Requirements Study (AORS) Raw Survey Results

Question
01

Which organization do you work for?

Answers 76 86%

Skips 12 14%

- Santa Clara County OES
- Harris County (TX) Homeland Security & Emergency Management
- NOAA/National Weather Service
- Clark County Emergency Management
- Calumet County, Wis.
- Galax Va Police
- Oklahoma Department of Emergency Management
- Maryland Emergency Management Agency
- City of Baytown, Texas Office of Emergency Management
- Iowa County
- Monroe County
- City of Galax Police Department
- Maryland Emergency Management Agency
- Indiana Department of Homeland Security
- AZ Department of Emergency & Military Affairs
- Central Ohio Trauma System
- Harris County (TX) Homeland Security & Emergency Management
- San Francisco Department of Emergency Management
- New York City Emergency Management
- Office of Emergency Management
- KCMO Office of Emergency Mgmt
- Minnesota Department of Public Safety - Division of Homeland Security and Emergency Management
- Santa Clara County
- Ingham County Emergency Management
- Hydroelectric Group
- City of Los Angeles
- Wisconsin Emergency Management
- County Emergency Management
- State of Iowa
- Millcreek Township Office of Emergency Management
- Saline County (Nebraska) Emergency Management
- Center for Disaster Research & Education (CDRE) at Millersville University
- Rail
- Lexington Kentucky Urban County Government Division of Emergency Management
- Lexington Fayette Urban County Government Division of Emergency Management
- Clark Regional Emergency Services Agency
- NOAA National Weather Service
- Anoka County Emergency Management
- Fairfield County Emergency Management and Homeland Security
- San Mateo County Office of Emergency Services
- Mississippi Emergency Management Agency

- Ventura County Sheriff's Office - Office of Emergency Services
- New Jersey State Office of Emergency Management
- Ohio AMBER Alert Advisory Committee
- County of San Diego Office of Emergency Services
- San Luis Obispo County (CA) Office of Emergency Services (County OES)
- Mono County Sheriff's Department, California
- City of Los Angeles Emergency Management Department
- Jay EMA
- Columbus Ohio Division of Police
- City of Alexandria Virginia
- Emergency management
- Outagamie County
- Emergency Management
- Penn State Office of Emergency Management
- County government
- City of Los Angeles
- Alabama Emergency Management Agency
- Mono County Sheriff's Office
- Salem County, NJ OEM
- Warren County New Jersey Department of Public Safety
- Washington State Military Department
- St. Clair County Homeland security - Emergency Management
- FEMA IPAWS Program Management Office
- Dane County
- Florida Division of Emergency Management
- Mississippi Emergency Management Agency
- Los Angeles Police Department
- Arkansas Department of Emergency Management
- Mississippi Emergency Management Agency
- WV DHSEM
- Mountain View Fire Department
- Mobile County Emergency Management Agency
- NA
- Maryland Emergency Management Agency

Question 02

What is your role within this organization?

Answers 76 86% Skips 12 14%

- Public Affairs Manager
- Sr. Emergency Services Coordinator
- Emergency Public Information Planner
- Emerging Dissemination Technologies Lead
- Director
- Emergency Management Director
- Chief
- EOC Manager
- Regional Liaison Officer
- Emergency Management Coordinator
- Emergency Management Director
- Emergency Management Coordinator
- Deputy Chief

- Manager of the 24/7 Watch Office that issues EAS messages
- […] Point Coordinator
- Operations Section; Human Services Group Supervisor
- Regional Healthcare Preparedness Coordinator
- Emergency Public Information Planner
- Response Coordinator
- Public Warning Specialist
- Deputy Coordinator
- Tech Coordinator
- Communications and Warning Officer
- Sr. Emergency Planning Coordinator, focused on EOC technology
- Deputy Emergency Manager
- Emergency Manager
- Manager in Emergency Management Department
- Office Operations Associate
- Director
- Homeland Security and Emergency Management
- Emergency Management Coordinator
- Director
- Graduate Assistant
- Superintendent
- Emergency Systems Specialist
- Emergency Systems Specialist
- Emergency Manager Coordinator
- Manager of a weather forecast office (Meteorologist-in-charge)
- Emergency Management Specialist
- Director
- OES District Coordinator
- Individual Assistance Specialist
- Director
- State Police Communications Bureau / NJOEM ESF-2 Communications
- AMBER Alert Coordinator
- Senior Emergency Services Coordinator responsible for Alert and Warning
- Emergency Services Manager (oversee County OES)
- Jail / Communications Commander
- Public Health Planner
- Director
- Communications Bureau Commander (911 call center/PSAP)
- Emergency Management Coordinator
- Warning Manager
- Director
- Coordinator
- Planner
- Department Director
- Public Information Officer
- Executive / Field Services / Division Coordinator
- Patrol / Emergency Management
- OEM Coordinator
- Director
- Telecommunications Engineer
- Director
- Deputy Director

Question 03

How many years have you worked in the emergency services domain?

Answers 79 90% Skips 9 10%

	COUNT	PERCENT
10+ years	55	70%
5 - 9 years	16	20%
0 - 4 years	8	10%

Question 04

How frequently does your organization issue emergency alerts?

Answers 78 90% Skips 10 10%

	COUNT	PERCENT
On average 1 to 5 times a month	48	63%
My organization does not issue emergency alerts	18	24%
On average 5 to 10 times a month	6	8%
On average higher than 10 times a month	4	5%

Question 05

Does your organization use the WEA service?

Answers 78 89% Skips 10 11%

	COUNT	PERCENT
Yes	44	56%
No	29	37%
Not sure	5	6%

Question 06

If you are currently using the WEA service, what is your primary motivation for the adoption and use of WEA?

Answers 48 55% Skips 40 45%

- Ability to reach large % of population in given area.
- WEA usage is restricted to extreme emergency situations - evacuations, national security issues.
- Protection of life and property from weather related threats.
- High-impact events, but particularly significant weather events (destructive tornadoes).

- n/a
- We have access to WEA, but due to the County-Wide message distribution, don't use it notify City residents
- IPAWS
- Improved effectiveness of alert and warning delivery to the public
- Fast, broad alerting ability.
- We are currently capable of issuing WEA notifications; however, rarely does an opportunity present itself. The motivation to use WEA is to reach as many people in an affected area as possible without a prior registration process. Since tourism is such a large industry, a prior registration would leave out a large number of people within any specific area.
- For shelter-in-place and evacuation messages
- Unlikely we will ever use it, but we need it available. It still has too many bugs, but is the future.
- Radiological Program
- Accessing cellphones
- We realize that public mass notification is the key to alerting the general public of immediate life/safety situations.
- Last tool in the tool box to make emergency notifications to the public 'complete'......landlines are going away and we have a large tourist population on a regular basis.
- Allowing National Weather Service the ability to send WEA for other than weather events. Prior to WEA, NWS served as our backup to send out alerts that were not weather related.
- To notify the public of local emergencies such as chemical spills for shelter-in-place or evacuation procedures.
- To notify the public of local emergencies such as chemical spills for shelter-in-place or evacuation procedures.
- Have our FEMA COG, still working with third party vendor for options on issuing WEA alerts.
- It was seamless for my office. It was implemented at a national level.
- Being able to send alerts to members of the community. Will be WEA capable this week. Final stages of installing and setting up EAS Encoder
- I believe it to be the future of alerting. As wireless adoption grows the reach to citizens also grows.
- Direct and immediate notification of citizens.
- The decision was made by our Executive Staff.
- We will only use WEA for a few very specific scenarios. Such as Tsunami evacuation alerts.
- Although Television and Broadcast outlets have State purchased and maintained equipment to receive and retransmit EAS messages, there is not much broadcaster participation in the retransmitting of EAS messages.
- National Center for Missing & Exploited Children (NCMEC) sets off WEA as part of AMBER Alert secondary systems.
- WEA allows OES to provide mass emergency information to individuals in an affected area. The primary benefit is that is another tool that allows us to be able to reach the whole community. Specifically, individuals that have cell phones and are actually in an affected area.
- Widespread dissemination of urgent emergency information.
- We have a large tourist base.
- emergency public notification.
- The ability to reach people in emergency situations that may be directly impacted by the event
- urgent notification of emergency situation.
- N/A
- Timely emergency notifications
- Notifying our residents of any eminent or actual treats in our City.

- Emergency Notification
- Targeted notification
- FEMA IPAWS is the gateway between public safety authorities using WEA and the wireless providers. The IPAWS Program built and sustains the gateway IT infrastructure and provides administration and assistance to users of IPAWS that send emergency alerts including WEA.
- Geographically targeted capability to reach large numbers of people, almost instantaneously, using devices that they own and carry with them everywhere.
- The use of WEA is required under some of our Comprehensive Emergency Management Plans.
- WEA aligns well with the reality that increasingly people are carrying communication devices and disconnecting copper wire phones.
- State-wide notification of an impact such as I40 being closed due to flooding
- Civil Emergency Messages
- Providing emergency information to persons who are mobile, in homes and are transient and don't belong to a local notification service.
- We would use WEA to inform the public of Hazmat accidents/incidents, or other emergencies, usually at the request first responders. We have approval/access to WEA, but have not actually used it yet.
- Reach and no sign up required by public

07

If your organization does NOT use the WEA service, why not? Rank the following reasons in terms of their importance for NOT using the WEA service. (Drag the responses below to reorder from highest to lowest importance.

Answers 88 100% Skips 0 0%

Rank Choice	Reported Rank	Weighted Rank
1. We are not set up to use the WEA service	1	3.52
2. We are using alternative alert messaging services.	2	3.98
3. We are not trained in issuing WEA messages.	3	4.35
4. Alternative alert messaging services are better.	4	4.45
5. WEA messages cannot target precisely enough.	5	4.58
6. WEA messages are hard to create.	6	4.65
7. WEA messages cause public complaints about false positives	7	4.80
8. It is hard to get WEA messages right.	9	5.02

Question 08

How likely are you to adopt and use WEA if the minimum geographic area that you can define for the delivery of the alert is

Answers 67 76% Skips 21 24%

	VERY UNLIKELY	UNLIKELY	LIKELY	VERY LIKELY
as small as 1 square city block	11	3	15	37
as small as 1 neighborhood (10 city blocks by 10 city blocks)	8	4	23	31
1 square mile	6	5	27	26
as large as 10 square miles	11	10	23	19
as large as 100 square miles	27	13	9	17

Question 09

The current WEA service delivers ALL alerts to ALL active, compatible mobile devices within the alert area. How would your willingness to adopt and use WEA change if WEA included the ability for recipients to choose to receive alerts based on the type and content of the alert message beyond the current ability to opt out of certain categories of alerts.

Answers 65 74% Skips 23 26%

	COUNT	PERCENT
Increased	33	51%
Decreased	17	26%
Greatly increased	11	17%
Greatly decreased	4	6%

Question **10**

The current WEA service delivers ALL alerts to ALL active, compatible mobile devices within the alert area. How would your willingness to adopt and use WEA change if WEA included the ability for recipients to choose to receive alerts based on their context at the time of the alert? (e.g., opt out of AMBER Alerts between midnight and 0600, suppress alerts while driving, receive alerts for an emergency in the vicinity of their residence even if they are away from home.)

Answers 65 74% Skips 23 26%

	COUNT	PERCENT
Increased	35	54%
Decreased	18	28%
Greatly increased	9	14%
Greatly decreased	3	5%

Question **11**

How likely are you to adopt and use WEA if the WEA service provides...

Answers 63 72% Skips 25 28%

	VERY UNLIKELY	UNLIKELY	LIKELY	VERY LIKELY
... no information on who has received the alert	9	18	26	8
... information on who has received the alert	2	5	37	18
... information on who has received the alert, and their location	2	5	33	23
... information on who has received the alert, their location, and when the alert	2	6	28	27
... information on who has received the alert, their location, when the alert was received, and recipient sentiment (like/dislike) toward the alert	4	14	24	20

Question
12

In deciding to adopt and use WEA, how important are the following features in alert origination tools?

Answers 64 73% Skips 24 27%

	VERY UNIMPORTANT	UNIMPORTANT	IMPORTANT	VERY IMPORTANT
The ability to easily define a geographic area for alert distribution (e.g., draw polygons on a map)	2	0	8	53
The ability to easily complete the information required to issue a WEA alert in the Common Alerting Protocol (CAP) format	3	2	9	50
The ability to check alert message spelling and length	1	2	23	38
The ability to facilitate multiple levels of internal authorization and approval for sending a WEA message	4	12	30	18

Question
13

Recognizing the trade-off between message length and limits and cost of wireless broadcast technology, how effective do you believe the WEA service can be if the maximum message length is...

Answers 67 76% Skips 21 24%

	VERY INEFFECTIVE	INEFFECTIVE	EFFECTIVE	VERY EFFECTIVE
90 characters	3	21	36	5
280 characters	0	5	27	35
500 characters	5	16	18	27
1000 characters	20	18	12	16

Question

14

A WEA message is like an alarm bell. Once the alarm is sounded, the message fulfills its purpose.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Disagree	30	43%
Agree	26	37%
Strongly disagree	8	11%
Strongly agree	6	9%

Question

15

A WEA message should do more than just sounding an alarm bell. It should give the recipient enough actionable information.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Strongly agree	43	61%
Agree	26	37%
Disagree	1	1%
Strongly disagree	0	0%

Question

16

A WEA message should be followed up with regular status updates to inform the public about any significant developments in the underlying emergency.

Answers 68 77% Skips 20 23%

	COUNT	PERCENT
Agree	36	53%
Strongly agree	21	31%
Disagree	10	15%
Strongly disagree	1	1%

Question 17

A WEA message should be followed up with a closure update to inform the public when the emergency is over.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Agree	31	44%
Strongly agree	29	41%
Disagree	10	14%
Strongly disagree	0	0%

Question 18

A WEA message should be followed up with a success update to inform the public when actions of emergency personnel or private residents help address the emergency.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Disagree	35	50%
Agree	24	34%
Strongly disagree	7	10%
Strongly agree	4	6%

Question 19

The alert originator community must rethink the vision of the WEA service within the entire alert messaging landscape.

Answers 69 78% Skips 19 22%

	COUNT	PERCENT
Agree	42	61%
Disagree	15	22%
Strongly agree	10	14%
Strongly disagree	2	3%

Question
20

Public education regarding WEA has/will have a strong influence on the effectiveness of WEA.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Strongly agree	40	57%
Agree	27	39%
Disagree	2	3%
Strongly disagree	1	1%

Question
21

To date, education of the public regarding WEA has been sufficient.

Answers 69 78% Skips 19 22%

	COUNT	PERCENT
Disagree	40	58%
Strongly disagree	23	33%
Agree	4	6%
Strongly agree	2	3%

Question
22

Education of the alert originator community regarding WEA has/will have a strong influence on the effectiveness of WEA.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Agree	35	50%
Strongly agree	32	46%
Disagree	3	4%
Strongly disagree	0	0%

Question 23

To date, education of the alert originator community regarding WEA has been sufficient.

Answers 70 80% Skips 18 20%

	COUNT	PERCENT
Disagree	40	57%
Strongly disagree	16	23%
Agree	11	16%
Strongly agree	3	4%

Question 24

Alert messages should be enriched with maps to help recipients determine how close they are to an affected area, even if doing so complicates and delays the alert generation process and inhibits the willingness of cell carriers to participate in the WEA service.

Answers 71 81% Skips 17 19%

	COUNT	PERCENT
Disagree	41	58%
Agree	19	27%
Strongly disagree	7	10%
Strongly agree	4	6%

Question 25

The alerting process would be enhanced by more integration of WEA with social media platforms.

Answers 71 81% Skips 17 19%

	COUNT	PERCENT
Agree	43	61%
Strongly agree	21	30%
Disagree	5	7%
Strongly disagree	2	3%

Question
26

Alert messages should be enriched with photos to help recipients identify landmarks and people affected by an emergency, even if doing so complicates and delays the alert generation process and inhibits the willingness of cell carriers to participate in the WEA service.

Answers 71 81%   Skips 17 19%

	COUNT	PERCENT
Disagree	45	63%
Agree	14	20%
Strongly disagree	8	11%
Strongly agree	4	6%

Question
27

Alert messages should be enriched with links to external sources to allow the public to easily access more information about an emergency, even if doing so runs the risk of overloading the cellular communication network.

Answers 70 80%   Skips 18 20%

	COUNT	PERCENT
Agree	29	41%
Disagree	25	36%
Strongly agree	10	14%
Strongly disagree	6	9%

Question
28

Which single improvement to the WEA service would make a real difference in their adoption by alert originators?

Answers 51 58%   Skips 37 42%

- links or increase message length
- Uniform adoption across all agencies
- message length
- A map showing the threat area and the recipients location
- Consistency among wireless providers
- understanding of a consistent national use of the system to get people's attention in regards to a risk
- Increasing characters beyond the 90 limit AND allowing us to test WEA.
- Ability to discriminate messages down to 1 block square areas by polygons
- Improved message length, streamline process for approval to send messages
- Educating providers on how to use
- Explaining how powerful the tool is and make sure they understand how the public may react once they push send
- Ease of use
- ability to add links to additional information
- Test environments.
- Increase in character count (100-300)
- Easier polygon selection without have to buy entire notification system.
- More than 90 Characters, even at 140 you could put out a more effective message.
- Longer messaging

- More direct control over language in message by originating authorized agency.
- Being able to deliver to targeted geographic areas rather than the entire service area
- Public Outreach should be prioritized.
- Increased State/Hyper Local participation (Townships, Cities, Etc)
- Education and public acceptance
- Communication
- Confidence in sending an alert. Allow agencies who have a FEMA issued COG the freedom of sending "required monthly tests" that people could opt out of on their mobile device.
- Better understanding of what type of systems are needed to issue WEA
- Increase the character length above 90 characters
- Accuracy and small scale of the alert area.
- Simplicity.
- Longer message. Defined geographical area.
- Adopt a national standard for how wireless devices are identified for notification when a shape file is drawn. One model is to identify wireless devices when a tower is in the shape file and the other model is to identify the device when the tower has a sector that provides coverage into the location of the shape file.
- Truly understanding the broadcast area (tower coverage area) that will receive a WEA message.
- Able to specify specific geographical areas to alert.
- Social media connections that contain detailed information, pictures, maps, etc.
- additional carriers
- accessibility to affordable platform to push out WEA
- Clearly defining who can originate emergency messages and the definition of emergency messages
- Ease of use over local notification methods
- a more precise way to focus the message to a geographic area.
- GPS position of the receiver that is across jurisdiction and or state boundaries.
- Training
- More education for the agencies that may use the system.
- Public awareness education
- ability to test system locally
- Refining the alert location to block level, truly alerting only those that are at risk.
- Education of the alert originators. Some alert originators are scared to use the system.
- The ability to test annually in a high hazard area or participating community.
- Cost of system
- ability to polygon down to small areas.
- Additional Characters
- Being able to generate and send WEA easily and quickly under stressful conditions.

Question **29** Which single improvement to the WEA service would make a real difference in their acceptance by the public?

Answers 54 61% Skips 34 39%

- links or increase message length
- Ability to opt out, especially via type of alert and time of day
- More granularity in what types of alerts they receive.
- message length
- More PROACTIVE outreach by wireless industry
- Education about WEA service and uses
- understanding of risks and public risk based response actions
- Issuance of meaningful alerts (message and location) and required action(s).
- All Hazards warning system which targets neighborhood level warnings. Current County-wide or regional is far too large.
- Education
- Being able to OPT out of messages of certain types, or during certain times
- Education
- inclusion of maps
- Consistent Public Education
- Education of the service to the public to illustrate the severity of the incident
- More targeted alerts. Irrelevant alerts hurts [sic] us all.
- More than 90 Characters, even at 140 you could put out a more effective message to tell them of the danger and to be able to put out a follow up message.
- Ability to customize alert types received
- Keeping issuance only for major catastrophes at the state and national level.
- Clear, short messages with direction on actions.
- Education about who can use it and what they can use it for
- Not being abused by sending out alert that could have been more affectively delivered via a different means.
- Additional character length
- Ease of use
- Safety
- Thorough understanding of how it works any why it is used. Several people have asked how they would be made aware of an existing alert after entering into the 'hot cell tower zone'.
- More education
- Standards for receiving alerts. Zip code based, county based, cell coverage based.
- Education.
- Short, simple message with link(s) to external websites.
- More public outreach and education.
- There must be enough information to alert the public to take appropriate corrective and protective measures but you cannot overload them with information. Addition info, updates, conclusion can be processed by a Public Information Officer and may use websites, social media and the broadcast media.
- Providing actionable and specific information. 90 characters is too short.
- Increased word count, in order to provide more information.
- Ability of the recipient to select which alerts they receive and block alerts during specified time frames
- education and better local control of messaging
- Education
- Geo location, not getting messages for emergencies that will not impact you.
- not sure whether this achievable
- use it only when absolutely needed (based on thresholds)
- Do not OVER WARN and become annoying.
- Text Length
- More education
- Public awareness education
- only imminent life threatening use
- more opt-in/opt-out categories (and education)

- More information on the system and how it works.
- Opt out of Amber alerts
- Education of the public. The public does not understand what they are.
- Send meaningful messages more frequently but to highly specific areas and more quickly than the news.
- outreach advertising
- providing actionable information
- Easily understandable information
- Education of the public about the real benefits of WEA, and how to use the information they get over WEA to increase their safety and awareness. Use of WEA in drills and exercises to demonstrate this to the public may also help.

Question **30** Enter any comments you might have about the survey, the WEA service, or about emergency alert messaging here.

Answers 25 28% Skips 63 72%

- The information provided on the WEA service by government agencies and/or cell providers has been virtually non-existent. I'm a professional in the field of crisis communications and have no idea how I could even hope to issue one. This would be a great tool and I hope to learn more about it soon.
- Improvements to WEA are, of course, welcome. However, it should be remembered that the purpose of WEA is to provide fast, actionable information. The more features allowed, the less useful it becomes as an "alarm bell."
- - Question 8 asks about WEA minimum geo-targeting resolution. I think it is a good question, but needs clarity. For example, does 10 square miles mean that the geotargeting could be off by as much as 10 square miles? - Question 9 states says, "The current WEA service delivers ALL alerts to ALL active, compatible mobile devices within the alert area. How would your willingness to adopt and use WEA change if WEA included the ability for recipients to choose to receive alerts based on the type and content of the alert message? (E.g., opt out of AMBER Alerts, opt out of weather alerts.)" However, WEA currently allows you to separately opt out of weather (celled Emergency Alerts on iPhone and severe/extreme on Android) and Amber Alerts. - There is an error in questions 11. The response in the right hand column of "Very unlikely" is supposed to be "Very Likely". Currently, you have "Very unlikely" on both the left and right. - Overall, the survey would benefit from a "neither agree nor disagree" response option for many of the questions.
- WEA is a great concept and while Oklahoma has not used it for an actual event, we have a plan in-place to do so. There are numerous limitations to WEA, many are covered in this survey. For use the 90 character limitation is an issue along with not being able to test the system, whereas EAS is required to be tested monthly. There is little training or guidance, only stated requirements to use the service. Public outreach is nonexistent. I would like to obtain best practices from other originators, mainly use-case reviews, after action reviews, etc. Public feedback would be nice but also requires a more involved research study. This survey instrument missed a coding area that I do feel is important. On many of the questions I had to select an answer instead of having the option to say "no opinion." Also, as noted at the beginning, my agency uses EAS/WEA sparingly. As of today we have not used WEA for an actual event, but may do so in the future.

- While there are other warning services providers (Code Red, Blackboard, etc.) WEA has credibility with public. It has stagnated however, and needs greater levels of warning discrimination, integration with social media and ability to provide actionable information along with the warning message. Web Links, useable by smart phones, would be ideal!

- Technically, we have the ability to send WEA messages; however, based upon the inability to test the system, I have very low confidence the message will go out in a specified amount of time to be effective. Additionally, with such a powerful tool, better education to alert originators is an absolute must and should be developed and required to issue alerts.
- Adding photos, maps, and links to WEAs may be beneficial to the public, but not if the messaging is delayed as a result. The purpose of the WEA is to disseminate actionable and timely messages to the public during a major disaster. If a separate public messaging system is already established (Notify NYC), photos, maps, and links can be added to public messages distributed through this system.
- have utilized more effective alternative means for alerts to date. Citizens may become confused or indifferent when there are too many methods for receiving alerts.
- Sounds like some good ideas. We need a better interface for JUST WEA that doesn't require 10's of thousands of dollars for full notification system. It has to be QUICK and SIMPLE in an emergency.
- WEA is a great addition to the toolbox for getting alerts sent out immediately for hazards that will most likely impact lives and property. I would recommend making this a mandatory for all wireless carriers to incorporate WEA into their networks.
- Need to learn more
- While trying to gain experience and confidence in sending an alert, we need emergency alert systems that work reliably. Several of our tests have failed while using different systems. We never know which one is really going to work when we send a message. In a live event, we have to have 2 backup systems in case the primary system fails.
- Education of the general public should be on the same scale as EBS/EAS. Testing can be done silently as to not become a nuisance. The ability to pinpoint location to an area as small as possible will also increase the effectiveness of the system. Amber Alerts also could be separated into another category of alerting. This will lessen the alerting fatigue seen in our State.
- The State of New Jersey does not currently issue many WEA messages. Our vendor is beta testing the polygon shape tool and at this time we use a county FIPS codes for EAS and WEA's. Small incidents would go countywide and we do not want the public to receive message that does not pertain to them. Don't issue rad and ignores. As the polygon tool is deployed I see and increase in use by County and even Municipal entities. Notifications regarding school lockdowns for suspicious persons, localized flooding, etc. Sergeant First Class John Doe XX State Police XX State Office of Emergency Management [ph XXX-XXX-XXX]
- In Ohio, we wanted the following options for a WEA message for AMBER alerts with and without vehicle information: With vehicle: "AMBER Alert, City, Ohio - Green Ford 4 door OH Reg ABC1234" Without vehicle" "AMBER Alert, City, Ohio - Tune to Local media" Even better would be able to place a website or phone number for the public to get further details!
- Since adding a URL to WEA messages increases the risk of potentially overloading the cell towers, a WEA Portal could be developed where alert originators could make a simplified webpage for the event (text only and only a couple kbs). This minor and potentially preferred traffic shouldn't overwhelm towers. Alert originators could place additional information on these specific webpages and provide this URL in the WEA message.
- This survey did not include a "neutral" answer. It was only agree or disagree. The survey started out asking questions as if I was considering implementing WEA. There should have been two surveys, one if you already use WEA and one if you do not.
- I know very little about WEA service. The remainder of the personnel in our PSAP are also minimally informed. As a major metropolitan area, it is obvious WEA could be valuable, but our use of WEA is nearly non-existent by our agency.
- WEA was a huge step forward for the alerting community. The ability to message unregistered cellular devices is essential to saving lives. WEA was the only way we were able to change behavior in a near catastrophic event shortly after WEA

implementation, the public did not react to text messaging, even on scene public safety personnel to take shelter but when the WEA alerts hit their phones they listened and took action. It changed behavior and worked well. Obviously improvements can always be made but please resist the urge to allow more opt out options and any quiet times!

- Good product but if there was a way to make it more precise so residents in areas not affected by the emergency are not receiving the message.
- Ask if and how you WEA to provide or receive emergency info. Drop down on state you live. Ask if you are aware of Public Safety's use in your local, county and state agencies. Ask if you think it should be used by educational institutions for All Hazards including Active Shooter. Poll participants and ask if they would be willing to complete future surveys and be a Beta User for testing. Ask if there should be certain cases where WEA Should NOT be used and explain why not.
- Q28 has a close second most improvement - increase the message length. The current message length is too short to do much more than sound an alarm. Most people need more information about an emergency situation before they will take action. Either a longer message or the ability to embed links is needed.
- I would like to see more education for alert originators to improve the knowledge of the agencies sending the alerts. I would also like to see a public awareness campaign centered around the alerts to educate the public why government is sending alerts.
- WEA is a vital tool in the overall toolbox of emergency managers. If used properly there is no better way to reach the public where they are.
- Currently we are authorized to use WEA, but have not yet had the need. We test our local warning systems, to include our Outdoor Siren Warning System, each month. If there were some way to include WEA in our monthly tests of all our warning systems, so that alert originators could develop and maintain a high level or proficiency in generating effective WEA alerts, and including them with other warnings systems, I do think this will increase the WEA success in a real emergency. I do very much believe in regular testing and training. I realize that sending a test WEA message to the general public may or may not be a good idea, as it may lead to warning fatigue, but at least the public will know what they are, and this is important. Here I am thinking of the tests the public sees on their TV of the EAS system. It may be that if special WEA test groups were available, such as to local first responders, or select agency heads or members, WEA tests could be sent to them only. It may be that this feature could also be used to notify this special group in certain emergency situations where the general public need not be notified. Also, it may be that if there is a way to have a special, unique, WEA ringtone, it may help increase the level of awareness. Thank you!

Question 31

If you would like to receive a summary of the survey results, please enter your email address below:

Answers 50 57% Skips 38 43%

Answers omitted.

Appendix B1. Emergency Scenarios Used in Experiment 3

Experiment 3 was divided into five scenarios. Each scenario was designed to test how different factors, such as complexity, number of parallel incidents and rate of change, would impact the subjects' awareness of the situation at hand. The scenarios are described below. They were presented in the given order.

Scenario 1: Earthquake with Plume

Duration: 27 hours *Number of Alerts: 5* *Number of parallel incidents: 2*

This scenario involves an earthquake warning followed by an update that tells the subjects to stay indoors due to the heavy damage caused by the earthquake. As a consequence of the earthquake, a local nuclear lab is affected and a radioactive plume is detected. This leads to an evacuation order. Simultaneously, a highway overpass collapses. Later in the day, the winds shift directions leading to a change in action, canceling the evacuation and directing everyone in the area to use a fallout shelter or stay underground.

This scenario was designed to interleave alerts of different incidents with unexpected changes to the information provided, but maintaining the overall instructions (what to do). Relatively few alerts were for the allotted time to give the subjects time to read and assimilate the information contained in the alerts.

Scenario 2: Random Alerts

Duration: 20 hours *Number of Alerts: 10* *Number of parallel incidents: 5*

The focus of this scenario was to increase the rate of change, but not complexity. This was achieved by creating and interleaving multiple simple, unrelated incidents. The scenario starts with a water supply pipe being ruptured and suspicion of water contamination. This incident is followed by a riot and a fire. The suspicion of water contamination is confirmed and proper instructions given. A missing child incident (—America's Missing: Broadcast Emergency Response—an Amber Alert) follows this update. Finally, an update is issued on the fire alert. Immediately after, the scenario concludes with a traffic incident.

The nature, or type, of the alerts issued, as well as the instructions provided with each alert, change frequently. However, this scenario involves very little change to the immediacy of the actions conveyed in the alerts.

Scenario 3: Severe Weather

Duration: 50 hours *Number of Alerts: 18* *Number of parallel incidents: 2*

This scenario revolves around an evolving weather emergency. It was designed to increase the complexity of an incident with multiple changes to the actions to be taken. It starts with an alert of heavy rain and wind. Severe weather leads to fallen power lines and flooded roads. The incident escalates to a widespread power outage, which in turn leads to opportunistic looting during the blackout. The weather gets worse, and the public is warned of flash floods and hail. Finally, civil unrest ensues and a curfew is imposed.

Throughout this scenario, the type of the emergency does not change much (almost all alerts were weather related). The scenario was played out over a relatively long period of time to evaluate the evolution of the subjects' situational awareness through the scenario. Multiple updates were issued in quick succession, but the polls were delayed by 10 to 15 hours to assess to what extent these quick changes were committed to memory.

Scenario 4: Alien Catastrophe

Duration: 49 hours *Number of Alerts: 25* *Number of parallel incidents: 5*

This was the most complex scenario used in the trial. We wanted to test an extreme case with intertwined incidents. To circumvent suspension of disbelief and engage the subjects, we deliberately designed the scenario to be fantastic and gave it a game-like nature. In this scenario, we upped not only complexity, but also rate of change; the nature and status of the underlying incidents were frequently updated.

The scenario involves multiple complex incidents, leading to multiple actions to take at any particular point in time, emulating an extreme case where the user needs to be aware of the entire situation to follow the storyline and know what to do.

The scenario starts with a satellite crashing in the middle of an intersection, followed by a seemingly unrelated virus infection in the same area. Traffic jams sprout throughout the city and the virus spreads to the outer communities. Violent crimes are reported and multiple fires erupt. These incidents lead to civil unrest, with multiple hospitals being mobbed. A few hours later, all hospitals reach full capacity; the virus has spread out to the entire county. The governor orders evacuations. During evacuations, multiple fires, riots and a bridge collapsing diverts the evacuations to different directions. In the middle of the second day, the virus is discovered to be of alien origin and highly contagious, which leads to a presidential alert mandating that the area be quarantined. All roads are subsequently blocked. By the end of the day, the World Health Organization discovers a treatment and informs everyone. However, following this announcement, supermarkets are mobbed by citizens who attempt to obtain the treatment before supplies run out. Martial law is imposed. The scenario ends with various closure updates terminating all open incidents.

Scenario 5: Bad Weather

Duration: 10 hours *Number of Alerts: 2* *Number of parallel incidents: 1*

The last scenario was designed to test the simplest case, with little complexity or change in the information provided. It involved only a couple of alerts. The goal was to test if the digest view feature had any impact under simple emergency circumstances.

The scenario starts with a severe thunderstorm and recommends that the citizens avoid driving. Strong winds pick up later in the day with a recommendation to stay at home.

APPENDIX B2. PUBLIC USABILITY TRIALS (PUT) DATA ANALYSIS APPROACH

This appendix explains the statistical techniques and approach used in analyzing the data for Experiments 1, 2 and 3 during the PUT phase of the work. It also provides the rationale for the approach and techniques.

The majority of the trial data collected were responses to multiple-choice feedback questions and polls issued after alerts, in-between alerts or at the end of an experiment. These questions and polls were intended to measure a set of outcome constructs, or dependent variables. The outcome construct levels represented in the questions and polls were of nominal scale (categorical), making it unsound to employ an analysis approach based on differences in mean values. The categorical, survey-like nature of the data made it most amenable to an analysis based on frequencies. The independent variables response rate and response delay were the only exceptions to this.

When testing response differences between two groups with respect to a tested factor (a control group and a factor group), we used the *Chi-square independence test* (Siegel 1994) with the null hypothesis that the tested factor is independent of the response distribution. Thus the null hypothesis states that the resulting outcome construct, or response variable, levels are independent of the value of the input variable, group affiliation determined by the presence or absence of a tested factor.

The Chi-square independence test is a non-parametric test in that it does *not* rely on any prior distributional assumptions about the data. It compares the response variable's expected conditional frequency distribution inferred from the data to the actual, observed conditional frequency distribution. In such data, any deviations from the expected distribution should follow a Chi-square distribution with the proper degree of freedom. If according to the underlying Chi-square distribution, the differences between expected and observed response level frequencies could not be simply due to chance with respect a selected confidence level, the null hypothesis is rejected and the alternative hypothesis that the response depends on the input is accepted.

When analyzing responses about overall impressions that do not depend on the presence or absence of an alert feature or characteristic, we used the *Chi-square goodness of fit test* (Siegel 1994). The goodness of fit is statistically similar to the independence test, except that the observed frequencies are not conditional on any input variable, and therefore the expected frequencies of the response levels are explicitly defined rather than inferred from the data. The test evaluates the null the hypothesis that the observed unconditional frequencies match the expected unconditional frequencies (that is whether the observed frequencies are a sufficiently good fit to the expected frequencies), and that any differences are likely to be due to chance with respect to a selected confidence level. We adopted a uniform distribution for the expected frequencies of the response levels, rejecting the null hypothesis when the differences could not be due to chance according to the underlying Chi-square distribution at the selected confidence level. The alternative hypothesis was that the responses were biased toward a specific set of response levels. The uniform distribution was a reasonable choice given that we did not have any prior knowledge or expectations regarding the frequencies of the response levels; we assumed that if the subjects were on average indifferent about a probed issue, their responses would be distributed evenly among the presented answer choices.

In all Chi-square tests, the selected confidence level was 95 percent. This means the null hypothesis was rejected, and the alternative hypothesis accepted, if there was less than a 5 percent chance that the differences between the observed and expected frequencies could be due to purely random effects, or chance. Expressed in statistical language, the chosen *alpha level*, or *statistical significance* was 5 percent, or 0.05, which is standard rule of thumb in hypothesis testing. Thus only results below this alpha level were considered statistically significant enough for hypothesis testing. Results at this alpha level or below are considered significant. As per the Chi-square test, statistical significance was evaluated by comparing the *Chi-square statistic* of an analyzed outcome to the corresponding threshold value at the desired alpha level of the underlying Chi-square distribution. The Chi-square statistic is the sum of the squares of the relative deviations from the expected frequencies.

In Chi-square independence tests, the tested factor (the input variable) always had two levels represented by the presence or absence of an enhancement feature or alert characteristic. There could be two or more categorical response levels (such as "Yes," "Unsure" and "No," or "Correct," "Partially Correct" and "Incorrect"). Here, the number of levels of the input and response variables determines the dimensions of the underlying contingency table (Zaiontz 2015) and the degree of freedom (df) of the underlying Chi-square distribution. In Chi-square goodness of fit tests, the number of response levels (categories) alone determines the degree of freedom of the underlying Chi-square distribution since the contingency table always has a single column. In all tests in the analysis, the df values were either 2 or 3. Correspondingly, the values of the threshold Chi-square statistic at an alpha level of 0.05 were 5.991 and 7.815, for df = 2 and df =3, respectively.

Statistical significance is only one component of hypothesis testing, related to the odds of an effect being real. Equally important is the magnitude of the effect, or *effect size* (Cohen 1992). Effect size answers the question: how pronounced is the effect in the response? For the independence test, effect size quantifies the magnitude of the differences in responses when a tested factor is present and absent. For the goodness of fit test, it quantifies the deviation from the expected distribution. The larger the effect size, the higher the tested construct's potential impact in practice, provided that the effect is also statistical significant, that is likely to be real rather than due to change.

In all Chi-square tests, we measured effect size in two different ways: (1) using *Cramer's V* statistic for independence tests and the *Phi* statistic for goodness of fit tests; and (2*) odds ratio* for both kinds of tests (Zaiontz 2015). Cramer's V and Phi represent theoretical measures, with standard rules for interpreting them, whereas the odds ratio is a practical measure, whose interpretation is more context-dependent. The odds ratio is reported only for results that are statistically significant at the chosen alpha level.

Phi and Cramer's V are calculated as follows:

$$\text{Phi} = \sqrt{\frac{x^2}{n}},\ \text{Cramer's V} = \sqrt{\frac{x^2}{n \cdot (k-1)}},$$

where x^2 is the Chi-square statistic, n is the number of observations, and k is smaller of the number of levels of the input variable (number of columns in the contingency table) and the number of levels of the response variables (number of row in the contingency table). The value of *k* was 2 in all independence tests; therefore, Phi and Cramer's V coincided in all cases. We interpreted theoretical effect sizes using the following, generally accepted rules of thumb given a constant value of *k* = 2 (Zaizontz, 2015; Cohen, 1992):

- Very small: Phi or Cramer's V smaller than 0.1
- Small: Phi or Cramer's V larger than or equal to 0.1 and less than 0.3
- Medium: Phi or Cramer's V larger than or equal to 0.3 and less than 0.5
- Large: Phi or Cramer's V larger than or equal to 0.5

An effect size between .2 and .3 as measured by Cramer's V or Phi is considered normal in studies dealing with human behavior where outcomes might be affected by multiple uncontrolled factors (Cohen 1992).

The odds of a positive response for a given group is defined as the ratio of the frequency of a positive response (for example, corresponding to the level "Yes" in some feedback responses) in that group to the frequency of a negative response (for example, corresponding to the levels "No" and "Unsure" together in some feedback responses) in that group. In all tested factors except the Situation Digest view, the mid-level responses are combined with the most negative level in calculating the odds ratio. In evaluating situational awareness with the Situation Digest view, the middle response level (Partially Incorrect) is ignored in the calculation of the odds ratio and only the absolute positive level (Correct) and absolute negative level (Wrong) are considered. Based on the scheme adopted, the odds ratio for a tested factor (an alert feature or characteristic) is calculated as the odds of a positive response when the tested factor is present (the feature group or characteristic group) to the odds of a positive response when the tested factor is absent (the control group). Thus the closer the odds ratio to 1, the smaller the effect size is. The interpretation of the odds ratio depends on the tested factor and its context.

Through Appendices B3-B6, we report all test results using tables that follow these conventions:

- H_0 denotes the null hypothesis. H_1 denotes the alternative hypothesis (accepted if H_0 is rejected).
- When analyzing a factor (an alert enhancement feature or alert characteristic) in an independence test, the "0" column signifies the absence of the factor (the control group) and the "1" column signifies the presence of the factor (the feature or characteristic group).
- Both absolute frequencies and relative frequencies (percentages) are reported for each applicable outcome (response) level of each outcome construct. For absolute frequencies, the "Tot." column reports the sums across underlying columns and rows. For relative frequencies, the "Tot." column reports the aggregated (or factor-independent) percentages, not the sum of percentages across underlying rows or columns.
- The value of the Chi-square statistic is reported under the column "Chi-Sq."
- If H_0 is rejected (and therefore H_1 is accepted) at the selected alpha level (the Chi-square value exceeds the threshold Chi-square value for that alpha level and applicable degree of freedom), the corresponding row under the "Rej. H_0?" column has a value "**Yes**" typeset in bold. Otherwise, the cell has a value "No".
- Under the left portion of the "Effect Size" column, we report either Phi or Cramer's V (denoted by C.V). If Phi replaces Cramer's V for certain rows, this is indicated in brackets in the corresponding cell.
- Odds ratios (denoted by O.R.) are reported in independence and goodness of fit tests only if the results are statistically significant. In goodness of fit tests for constructs related to overall impressions, odds ratio is omitted.

APPENDIX B3. PUBLIC USABILITY TRIALS (PUT) DATA TABLES AND STATISTICAL ANALYSIS — ALERT-BASED ENHANCEMENT FEATURES

B3.1. Long Message

Table B3.1. Experiment 1 — Analysis of Long Messages vs. Short Messages

Experiment 1	**Tested Factor: Enhancement Feature - Long Message**										
H_0: Measured level of outcome construct is independent of message length. H_1: Long messages improve the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	19	3	22	8	4	7	1.9	No	0.08	1.18
	Partially	16	7	23	7	9	7				
	Yes	198	67	265	85	87	85				
	Total	233	77	310	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	71	23	94	31	31	31	0.0	No	0.01	1.02
	Unsure	33	10	43	14	14	14				
	Yes	127	41	168	55	55	55				
	Total	231	74	305	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	162	62	224	71	83	74	4.7	No	0.12	0.70
	Somewhat	25	3	28	11	4	9				
	Yes	41	10	51	18	13	17				
	Total	228	75	303	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	77	17	94	33	24	31	5.3	No	0.13	0.96
	Unsure	31	17	48	13	24	16				
	Yes	122	37	159	53	52	53				
	Total	230	71	301	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	71	23	94	31	31	31	0.6	No	0.04	1.15
	Unsure	29	7	36	13	9	12				
	Yes	131	45	176	57	60	58				
	Total	231	75	306	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	77	17	94	34	24	32	2.8	No	0.10	1.36
	Unsure	28	11	39	12	15	13				
	Yes	121	44	165	54	61	55				
	Total	226	72	298	100	100	100				

Table B3.2. Experiment 2 — Analysis of Long Messages vs. Short Messages

	Experiment 2	**Tested Factor: Enhancement Feature - Long Message**									
H_0: Measured level of outcome construct is independent of message length. H_1: Long messages improve the measured level of outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding:	**No**	8	6	14	4	3	4	2.10	No	0.07	0.72
Did you understand this alert message?	**Partially**	8	13	21	4	7	6				
	Yes	183	156	339	92	89	91				
	Total	199	175	374	100	100	100				
Relevance:	**No**	58	34	92	20	25		7.28	**Yes**	0.14	1.76
In a real emergency, would this alert be relevant to you given your current situation and location?	**Not Sure**	36	25	61	18	14	16				
	Yes	104	115	219	53	66	59				
	Total	198	174	372	100	100	100				
Annoyance:	**No**	111	123	234	58	72	65	8.81	**Yes**	0.16	0.49
In a real emergency, would this alert annoy you?	**Somewhat**	34	24	58	18	14	16				
	Yes	46	23	69	24	14	19				
	Total	191	169	360	100	100	100				
Actionability:	**No**	67	37	104	35	22	29	9.25	**Yes**	0.16	1.81
In a real emergency, would this alert prompt you to take action?	**Not Sure**	39	32	71	20	19	20				
	Yes	85	100	185	45	59	51				
	Total	191	169	360	100	100	100				
Milling Behavior:	**No**	58	30	88	29	17	24	8.02	**Yes**	0.15	1.73
In a real emergency, would this alert prompt you to seek further information?	**Not Sure**	29	24	53	15	14	14				
	Yes	112	120	232	56	69	62				
	Total	199	174	373	100	100	100				
Adequacy:	**No**	40	32	72	22	20	21	0.45	No	0.04	1.04
Did this alert contain enough information?	**Not Sure**	11	12	23	6	7	7				
	Yes	131	118	249	72	73	72				
	Total	182	162	344	100	100	100				

Table B3.3. Experiments 1 & 2 Combined — Analysis of Long Messages vs. Short Messages

Experiment 1 and 2 Pooled	Tested Factor: Enhancement Feature - Long Message										
H_0: Measured level of outcome construct is independent of message length. H_1: Long messages improve the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	27	9	36	6	4	5	3.57	No	0.07	1.03
	Partially	24	20	44	6	8	6				
	Yes	381	223	604	88	88	88				
	Total	432	252	684	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	129	57	186	30	23	27	5.52	No	0.09	1.45
	Not Sure	69	35	104	16	14	15				
	Yes	231	156	387	54	63	57				
	Total	429	248	677	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	273	185	458	65	76	69	8.07	**Yes**	0.11	0.59
	Somewhat	59	27	86	14	11	13				
	Yes	87	33	120	21	13	18				
	Total	419	245	664	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	144	54	198	34	23	30	10.05	**Yes**	0.12	1.38
	Not Sure	70	49	119	17	20	18				
	Yes	207	137	344	49	57	52				
	Total	421	240	661	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	129	53	182	30	21	27	7.09	**Yes**	0.10	1.51
	Not Sure	58	31	89	13	12	13				
	Yes	243	165	408	57	66	60				
	Total	430	249	679	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	117	49	166	29	21	26	4.74	No	0.09	1.39
	Not Sure	39	23	62	10	10	10				
	Yes	252	162	414	62	69	64				
	Total	408	234	642	100	100	100				

B3.2. High-Information Map

Table B3.4. Experiment 1 — Analysis of Alerts With and Without High-Information Maps

Experiment 1	**Tested Factor: Enhancement Feature - Map with Location**										
H_0: Measured level of outcome construct is independent of inclusion of a map showing geo-target and location. H_1: Alerts with maps showing geo-target and location improve the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	19	5	24	8	4	7	3.3	No	0.10	1.96
	Partially	16	5	21	7	4	6				
	Yes	198	111	309	85	92	87				
	Total	233	121	354	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	71	32	103	31	28	30	10.5	**Yes**	0.17	1.75
	Unsure	33	4	37	14	4	11				
	Yes	127	77	204	55	68	59				
	Total	231	113	344	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	162	91	253	71	76	73	1.4	No	0.06	0.87
	Somewhat	25	9	34	11	8	10				
	Yes	41	19	60	18	16	17				
	Total	228	119	347	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	77	36	113	33	32	33	0.2	No	0.02	1.02
	Unsure	31	17	48	13	15	14				
	Yes	122	61	183	53	54	53				
	Total	230	114	344	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	71	1	72	31	9	30	2.9	No	0.11	3.44
	Unsure	29	1	30	13	9	12				
	Yes	131	9	140	57	82	58				
	Total	231	11	242	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	77	2	79	34	18	13	2.6	No	0.10	1.04
	Unsure	28	3	31	12	33	55				
	Yes	121	6	127	54	27	54				
	Total	226	11	237	100	100	100				
H_0: Subjects find maps useful, do not find them useful, and are unsure in similar proportions. H_1: More subjects find maps useful than those who find them not useful or are unsure.											
Usefulness: *Was the map sent with the alert useful?*	**Didn't notic**		6	6		6	6	152.6	**Yes**	1.18 (Phi)	
	No		10	10		9	9				
	Unsure		7	7		6	6				
	Yes		86	86		79	79				
	Total		109	109		100	100				

(Usefulness: Levels "Didn't notice" and "Unsure" are merged in Chi-Sq test)

Table B3.5. Experiment 2 — Analysis of Alerts With and Without High-Information Maps

Experiment 2	**Tested Factor: Enhancement Feature - Map with Location**										
H_o: Measured level of outcome construct is independent of inclusion of a map showing geo-target and location. H_1: Alerts with maps showing geo-target and location improve the measured level of outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding:	**No**	7	7	14	3	3	3	0.34	No	0.03	1.19
Did you understand this alert message?	**Partially**	13	11	24	5	4	5				
	Yes	227	244	471	92	93	93				
	Total	247	262	509	100	100	100				
Relevance:	**No**	63	65	128	26	25	25	5.58	No	0.11	1.35
In a real emergency, would this alert be relevant to you given your current situation and location?	**Not Sure**	41	25	66	17	10	13				
	Yes	142	166	308	58	65	61				
	Total	246	256	502	100	100	100				
Annoyance:	**No**	148	161	309	61	62	62	0.07	No	0.01	0.99
In a real emergency, would this alert annoy you?	**Somewhat**	42	43	85	17	17	17				
	Yes	51	54	105	21	21	21				
	Total	241	258	499	100	100	100				
Actionability:	**No**	75	88	163	31	34	33	2.67	No	0.07	1.13
In a real emergency, would this alert prompt you to take action?	**Not Sure**	57	46	103	24	18	21				
	Yes	108	124	232	45	48	47				
	Total	240	258	498	100	100	100				
Milling Behavior:	**No**	56	76	132	23	29	26	2.58	No	0.07	0.81
In a real emergency, would this alert prompt you to seek further information?	**Not Sure**	42	42	84	17	16	17				
	Yes	148	144	292	60	55	57				
	Total	246	262	508	100	100	100				
Adequacy:	**No**	48	37	85	21	15	18	5.46	No	0.11	1.12
Did this alert contain enough information?	**Not Sure**	16	29	45	7	11	9				
	Yes	164	189	353	72	74	73				
	Total	228	255	483	100	100	100				
H_o: Subjects find maps useful, do not find them useful, and are unsure in similar proportions. H_1: More subjects find maps useful than those who find them not useful or are unsure.											
Usefulness:	**No**		28	28		11	11	306.2	**Yes**	1.08 (Phi)	
Was the map sent with the alert useful?	**Not Sure**		13	13		5	5				
	Yes		220	220		84	84				
	Total		261	261		100	100				

Table B3.6. Experiments 1 & 2 Combined — Analysis of Alerts With and Without High-Information Maps

Experiment 1 and 2 Pooled	**Tested Factor: Enhancement Feature - Map with Location**

H_o: Measured level of outcome construct is independent of inclusion of a map showing geo-target and location.
H_1: Alerts with maps showing gee-target and location imp rove the measured level of outcome constmct.

Outcome Construct: *Feedback Question*	Answer (Level)	Frequency 0	Frequency 1	Frequency Tot.	Percentage 0	Percentage 1	Percentage Tot.	Chi-Sq.	Rej. H0?	Effect Size C.V	Effect Size O.R.
Understanding: *Did you understand this alert message?*	**No**	26	12	38	5	3	4	4.35	No	0.07	1.64
	Partially	29	16	45	6	4	5				
	Yes	425	355	780	89	93	90				
	Total	480	383	863	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	134	97	231	28	26	27	13.34	**Yes**	0.13	1.49
	Not Sure	74	29	103	16	8	12				
	Yes	269	243	512	56	66	61				
	Total	477	369	846	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	310	252	562	66	67	66	0.06	No	0.01	0.98
	Somewhat	67	52	119	14	14	14				
	Yes	92	73	165	20	19	20				
	Total	469	377	846	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	152	124	276	32	33	33	0.46	No	0.02	1.03
	Not Sure	88	63	151	19	17	18				
	Yes	230	185	415	49	50	49				
	Total	470	372	842	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	127	77	204	27	28	27	0.43	No	0.02	0.90
	Not Sure	71	43	114	15	16	15				
	Yes	279	153	432	58	56	58				
	Total	477	273	750	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	125	39	164	28	15	23	15.86	**Yes**	0.15	1.63
	Not Sure	44	32	76	10	12	11				
	Yes	285	195	480	63	73	67				
	Total	454	266	720	100	100	100				

H_o: Subjects find maps useful, do not find them useful, and are unsure in similar proportions.
H_1: More subjects find maps useful than those who find them not useful or are unsure.

Outcome Construct: *Feedback Question*	Answer (Level)	Frequency 0	Frequency 1	Frequency Tot.	Percentage 0	Percentage 1	Percentage Tot.	Chi-Sq.	Rej. H0?	Effect Size C.V	Effect Size O.R.
Usefulness: *Was the map sent with the alert useful?*	**No**		38	38		10	10	840.6	**Yes**	1.51 (Phi)	
	Not Sure		26	26		7	7				
	Yes		306	306		63	63				
	Total		370	370		100	100				

Table B3.7. Overall Impressions Regarding High-Information Maps

Feature Preference — High-Information Map: Overall Impression (Clarity and Relevance)

H_0: Subjects are equally divided among their overall impression regarding the clarity and relevance of alerts delivered with a map showing the recipient's location.
H_1: More subjects find alerts delivered with a map showing the recipient's location clear and relevant than those who find them only somewhat clear and relevant or unclear and irrelevant.

Feedback Question: *Did inclusion of maps with some alerts increased their clarity and relevance?*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. H_0?	Effect Size Phi
Experiment 1	**No**			29.90	**Yes**	1.21
	Somewhat	2	10			
	Yes	18	90			
	Total	20	100			
Experiment 2	**No**			62.60	**Yes**	1.15
	Somewhat	6	13			
	Yes	41	87			
	Total	47	100			
Experiment 1 and 2 Pooled	**No**			91.73	**Yes**	1.17
	Somewhat	8	12			
	Yes	59	88			
	Total	67	100			

B3.3. Geo-targeting

Table B3.8. Experiment 1 — Analysis of Alerts With and Without Geo-Targeting

Experiment 1	**Tested Factor: Enhancement Feature - Geo-targeting**										
H_0: Measured level of outcome construct is independent of whether the alert was precisely gee-targeted or not. H_1: Alerts that were precisely gee-targeted improve the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	15	2	17	8	3	7	4.3	No	0.13	3.00
	Partially	14	2	16	8	3	6				
	Yes	157	65	222	84	94	87				
	Total	186	69	255	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	62	6	68	34	9	27	25.6	**Yes**	0.32	6.11
	Unsure	28	3	31	15	4	12				
	Yes	95	58	153	51	87	61				
	Total	185	67	252	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	128	53	181	70	79	73	4.1	No	0.13	1.05
	Somewhat	20	2	22	11	3	9				
	Yes	34	12	46	19	18	18				
	Total	182	67	249	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	61	16	77	33	24	31	2.9	No	0.11	1.63
	Unsure	25	8	33	14	12	13				
	Yes	97	44	141	53	65	56				
	Total	183	68	251	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	57	14	71	31	25	31	0.6	No	0.05	1.25
	Unsure	23	7	30	13	13	13				
	Yes	104	34	138	57	62	58				
	Total	184	55	239	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	68	9	77	38	16	33	9.0	**Yes**	0.20	2.21
	Unsure	21	8	29	12	15	12				
	Yes	90	38	128	50	69	55				
	Total	179	55	234	100	100	100				
H_0: Subjects outside the specific area targeted by an alert, but are in the general cosmopolitan area are equally interested, disinterested, and indifferent in receiving those alerts. H_1: More subjects outside the specific area targeted by an alert are disinterested in receiving those alerts than those who are interested or indifferent.											
Hindsight Relevance: *In a real emergency, would you have preferred to receive this alert?*	**No**	16		16		15	15	65.1	**Yes**	0.78 (Phi)	
	Unsure	16		16		15	15				
	Yes	75		75		70	70				
	Total	107		107		100	100				

Table B3.9. Experiment 2 — Analysis of Alerts With and Without Geo-Targeting

Experiment 2	**Tested Factor: Enhancement Feature - Geo-targeting**										
H_o: Measured level of outcome construct is independent of whether the alert was precisely gee-targeted or not. H_1: Alerts that were precisely geo-targeted improve the measured level of outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding:	**No**	5	1	6	3	3	3	0.44	No	0.05	1.52
Did you understand this alert message?	**Partially**	2		2	1	0	1				
	Yes	143	31	174	85	97	96				
	Total	150	32	182	100	100	100				
Relevance:	**No**	64	5	69	43	16	38	8.85	**Yes**	0.22	2.25
In a real emergency, would this alert be relevant to you given your current situation and location?	**Not Sure**	16	6	22	11	19	12				
	Yes	68	21	89	46	66	49				
	Total	148	32	180	100	100	100				
Annoyance:	**No**	84	21	105	58	68	59	1.49	No	0.09	0.53
In a real emergency, would this alert annoy you?	**Somewhat**	30	6	36	21	19	20				
	Yes	32	4	36	22	13	20				
	Total	146	31	177	100	100	100				
Actionability:	**No**	66	9	75	45	29	42	2.75	No	0.12	1.56
In a real emergency, would this alert prompt you to take action?	**Not Sure**	30	8	38	21	26	21				
	Yes	50	14	64	34	45	36				
	Total	146	31	177	100	100	100				
1Milling Behavior:	**No**	65	14	79	44	44	44	0.00	No	0.01	1.01
In a real emergency, would this alert prompt you to seek further information?	**Not Sure**	24	5	29	16	16	16				
	Yes	60	13	73	40	41	40				
	Total	149	32	181	100	100	100				
Adequacy:	**No**	17	4	21	12	14	13	1.51	No	0.09	1.24
Did this alert contain enough information?	**Not Sure**	7		7	5	0	4				
	Yes	116	24	140	83	86	83				
	Total	140	28	168	100	100	100				
H_o: Subjects outside the specific area targeted by an alert, but are in the general cosmopolitan area are equally interested, disinterested, and indifferent in receiving those alerts. H_1: More subjects outside the specific area targeted by an alert are disinterested in receiving those alerts than those who are interested or indifferent											
Hindsight Relevance:	**No**	42			28		28	41.73	**Yes**	0.48 (Phi)	
In a real emergency, would you have preferred to receive this alert?	**Not Sure**	22			15		15				
	Yes	85			57		57				
	Total	149			100		100				

Table B3.10. Experiments 1 & 2 Combined — Analysis of Alerts With and Without Geo-Targeting

Experiment 1and 2 Pooled	**Tested Factor: Enhancement Feature - Geo-targeting**										
H_0: Measured level of outcome construct is independent of whether the alert was precisely geo-targeted or not. H_1: Alerts that were precisely geo-targeted improve the measured level of outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	20	3	23	6	3	5	3.06	No	0.08	2.30
	Partially	16	2	18	5	2	4				
	Yes	300	96	396	89	95	91				
	Total	336	101	437	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	126	11	137	38	11	32	31.21	**Yes**	0.27	4.12
	Not Sure	44	9	53	13	9	12				
	Yes	163	79	242	49	80	56				
	Total	333	99	432	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	212	74	286	65	76	67	4.67	No	0.10	0.77
	Somewhat	50	8	58	15	8	14				
	Yes	66	16	82	20	16	19				
	Total	328	98	426	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	127	25	152	39	25	36	6.90	**Yes**	0.13	1.75
	Not Sure	55	16	71	17	16	17				
	Yes	147	58	205	45	59	48				
	Total	329	99	428	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	122	28	150	37	32	36	0.70	No	0.04	1.21
	Not Sure	47	12	59	14	14	14				
	Yes	164	47	211	49	54	50				
	Total	333	87	420	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	85	13	98	27	16	24	4.33	No	0.10	1.62
	Not Sure	28	8	36	9	10	9				
	Yes	206	62	268	65	75	67				
	Total	319	83	402	100	100	100				
Hindsight Relevance: *In a real emergency, would you have preferred to receive this alert?*	**No**	58		58	23		23	100.3	**Yes**	0.63 (Phi)	
	Not Sure	38		38	15		15				
	Yes	160		160	63		63				
	Total	256		256	100		100				

B3.4. External Link

Table B3.11 Experiments 1 — Analysis of Alerts With and Without External Links

Experiment 1	Tested Factor: Enhancement Feature - External Link										
H_o: Measured level of outcome construct is independent of inclusion of an external link or social media hashtag in the alert. H_1: Alerts with external links or social media hashtags improve the measured level of outcome construct.											
Outcome Construct:	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
Feedback Question		**0**	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	16	8	24	7	9	8	1.0	No	0.06	0.71
	Partially	15	8	23	7	9	8				
	Yes	188	69	257	86	81	85				
	Total	219	85	304	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	67	33	100	31	39	33	1.8	No	0.08	0.73
	Unsure	28	11	39	13	13	13				
	Yes	121	41	162	56	48	54				
	Total	216	85	301	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	149	52	201	70	61	67	2.2	No	0.09	1.52
	Somewhat	23	10	33	11	12	11				
	Yes	42	23	65	20	27	22				
	Total	214	85	299	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	79	29	108	37	35	36	2.1	No	0.08	0.84
	Unsure	27	16	43	13	19	14				
	Yes	110	39	149	51	46	50				
	Total	216	84	300	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	65	4	69	32	22	31	1.9	No	0.09	2.05
	Unsure	24	1	25	12	6	11				
	Yes	113	13	126	56	72	57				
	Total	202	18	220	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	70	6	76	36	33	35	1.9	No	0.09	0.70
	Unsure	22	4	26	11	22	12				
	Yes	105	8	113	53	44	53				
	Total	197	18	215	100	100	100				
H_o: Subjects find embedded external links or social media tags useful, do not find them useful, and are unsure in similar proportions.. H_1: More subjects find embedded external links or social media tags useful than those who find them not useful or are unsure.											
Usefulness: *Was the link embedded in this alert useful?*	**Didn't notic**		16	16		19	19	29.8	**Yes**	0.60	
	No		21	21		25	25				
	Unsure		6	6		7	7				
	Yes		41	41		49	49				
	Total		84	84		100	100				

Table B3.12 Experiments 2 — Analysis of Alerts With and Without External Links

Experiment 2	Tested Factor: Enhancement Feature - External Link										
H_o: Measured level of outcome construct is independent of inclusion of an external link or social media hashtag in the alert. H_1: Alerts with external links or social media hashtags improve the measured level of outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding:	**No**	6	4	10	8	5	6	0.96	No	0.08	1.76
Did you understand this alert message?	**Partially**	2	1	3	3	1	2				
	Yes	72	79	151	90	94	92				
	Total	80	84	164	100	100	100				
Relevance:	**No**	16	17	33	20	20	20	0.73	No	0.07	1.20
In a real emergency, would this alert be relevant to you given your current situation and location?	**Not Sure**	11	8	19	14	10	12				
	Yes	53	59	112	66	70	68				
	Total	80	84	164	100	100	100				
Annoyance:	**No**	60	55	115	76	66	71	5.52	No	0.18	0.94
In a real emergency, would this alert annoy you?	**Somewhat**	3	12	15	4	14	9				
	Yes	16	16	32	20	19	20				
	Total	79	83	162	100	100	100				
Actionability:	**No**	22	19	41	28	23	25	3.39	No	0.14	0.87
In a real emergency, would this alert prompt you to take action?	**Not Sure**	4	11	15	5	13	9				
	Yes	53	53	106	67	64	65				
	Total	79	83	162	100	100	100				
Milling Behavior:	**No**	20	21	41	25	25	25	0.58	No	0.06	0.85
In a real emergency, would this alert prompt you to seek further information?	**Not Sure**	6	9	15	8	11	9				
	Yes	54	53	107	68	64	66				
	Total	80	83	163	100	100	100				
Adequacy:	**No**	16	10	26	20	12	16	2.10	No	0.11	1.69
Did this alert contain enough information?	**Not Sure**	6	5	11	8	6	7				
	Yes	58	67	125	73	82	77				
	Total	80	82	162	100	100	100				
H_o: Subjects find embedded external links or social media tags useful, do not find them useful, and are unsure in similar proportions H_1: More subjects find embedded external links or social media tags useful than those who find them not useful or are unsure.											
Usefulness:	**No**		8	8		10	10	50.00	**Yes**	0.77	
Was the link embedded in this alert useful?	**Not Sure**		18	18		21	21				
	Yes		58	58		69	69				
	Total		84	84		100	100				

Table B3.13. Experiments 1 and 2 Combined — Analysis of Alerts With and Without External Links

Experiment 2	**Tested Factor: Enhancement Feature - Location Prediction**										
H_o: Measured level of outcome construct is independent of whether the alert was sent to users who are moving toward the alert area or not. H_1: Alerts that are targeted to recipients who are moving towards an alert area improve the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	22	12	34	7	7	7	0.04	No	0.01	1.06
	Partially	17	9	26	6	5	6				
	Yes	260	148	408	87	88	87				
	Total	299	169	468	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	83	50	133	28	30	29	0.41	No	0.03	1.02
	Not Sure	39	19	58	13	11	12				
	Yes	174	100	274	59	59	59				
	Total	296	169	465	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	209	107	316	71	64	69	3.33	No	0.08	1.22
	Somewhat	26	22	48	9	13	10				
	Yes	58	39	97	20	23	21				
	Total	293	168	461	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	101	48	149	34	29	32	3.72	No	0.09	0.99
	Not Sure	31	27	58	11	16	13				
	Yes	163	92	255	55	55	55				
	Total	295	167	462	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	85	25	110	30	25	29	1.25	No	0.06	1.30
	Not Sure	30	10	40	11	10	10				
	Yes	167	66	233	59	65	61				
	Total	282	101	383	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	86	16	102	31	16	27	9.28	**Yes**	0.16	2.10
	Not Sure	28	9	37	10	9	10				
	Yes	163	75	238	59	75	63				
	Total	277	100	377	100	100	100				
H_o: Subjects find embedded external links or social media tags useful, do not find them useful, and are unsure in similar proportions. H_1: More subjects find embedded external links or social media tags useful than those who find them not useful or are unsure											
Usefulness: *Was the link embedded in this alert useful?*	**No**		29	29		20	20	218.5	**Yes**	1.22	
	Not Sure		39	39		26	26				
	Yes		80	80		54	54				
	Total		148	148		100	100				

Table B3.14. Overall Impressions Regarding Alerts with External Links

Feature Preference — External Link: Overall Impression (Usefulness)						
H_0: Subjects are equally divided among their overall impression regarding the usefulness of alerts that contained links or social media tags. H_1: More subjects find alerts that contained links or social media links useful than those who are unsure or do not find them useful.						
Feedback Question: *Was the inclusion of links to external sources (URLs or social media tags) useful?*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. H_0?	Effect Size Phi
Experiment 1	**No**	4	20	45.50	Yes	1.51
	Unsure	1	5			
	Yes	15	75			
	Total	20	100			
Experiment 2	**No**	5	11	21.23	Yes	0.67
	Unsure	12	26			
	Yes	30	64			
	Total	47	100			
Experiment 1 and 2 Pooled	**No**	91	13	34.87	Yes	0.72
	Unsure	13	19			
	Yes	45	67			
	Total	67	100			

B3.5. Text-to-Speech

Table B3.15. Experiment 1 — Analysis of Alerts With and Without Text-to-Speech

Experiment 1	**Tested Factor: Enhancement Feature - Text-to-Speech**										
H_0: Measured level of outcome construct is independent of whether an alert is delivered with Text-to-Speech or not. H_1: Delivery of alerts with Text-to-Speech improves the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding:	**No**	23	8	31	12	4	8	15.2	**Yes**	0.20	3.08
Did you understand this alert message?	**Partially**	22	10	32	12	5	8				
	Yes	146	180	326	76	91	84				
	Total	191	198	389	100	100	100				
Relevance:	**No**	67	52	119	35	27	31	3.6	No	0.10	1.43
In a real emergency, would this alert be relevant to you given your current situation and location?	**Unsure**	27	27	54	14	14	14				
	Yes	96	115	211	51	59					
	Total	190	194	384	100	100	100				
Annoyance:	**No**	121	154	275	64	79	72	11.1	**Yes**	0.17	0.46
In a real emergency, would this alert annoy you?	**Somewhat**	21	15	36	11	8	9				
	Yes	46	25	71	24	13	19				
	Total	188	194	382	100	100	100				
Actionability:	**No**	71	45	116	38	23	31	10.0	**Yes**	0.16	1.68
In a real emergency, would this alert prompt you to take action?	**Unsure**	29	34	63	16	18	17				
	Yes	86	114	200	46	59	53				
	Total	186	193	379	100	100	100				
Milling Behavior:	**No**	49	52	101	35	27	30	3.0	No	0.09	1.46
In a real emergency, would this alert prompt you to seek further information?	**Unsure**	18	22	40	13	11	12				
	Yes	75	121	196	53	62	58				
	Total	142	195	337	100	100	100				
Adequacy:	**No**	39	61	100	28	32	30	0.6	No	0.04	0.95
Did this alert contain enough information?	**Unsure**	20	24	44	15	13	13				
	Yes	78	107	185	57	56	56				
	Total	137	192	329	100	100	100				

Table B3.16. Experiment 2 — Analysis of Alerts With and Without Text-to-Speech

Experiment 2 **Tested Factor: Enhancement Feature - Text-to-Speech**

H_0: Measured level of outcome construct is independent of whether an alert is delivered with Text-to-Speech or not.
H_1: Delivery of alerts with Text-to-Speech improves the measured level of outcome construct.

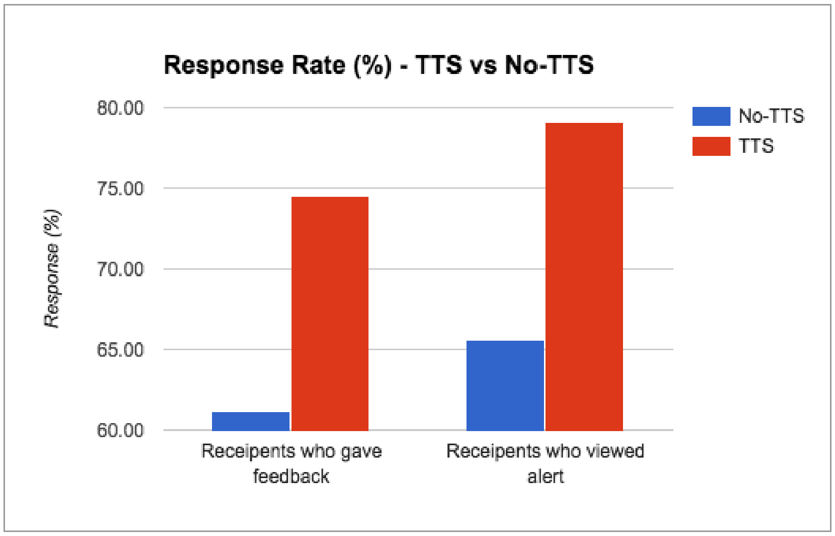
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	6	8	14	3	4	4	0.49	No	0.04	0.79
	Partially	9	14	23	5	6	6				
	Yes	168	195	363	92	90	91				
	Total	183	217	400	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	39	53	92	24	25	24	0.08	No	0.01	1.00
	Not Sure	29	36	65	18	17	17				
	Yes	97	127	224	59	59	59				
	Total	165	216	381	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	111	127	238	69	61	65	4.62	No	0.11	1.00
	Somewhat	18	40	58	11	19	16				
	Yes	31	40	71	19	19	19				
	Total	160	207	367	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	47	58	105	29	28	28	1.15	No	0.06	0.89
	Not Sure	28	46	74	18	22	20				
	Yes	85	105	190	53	50	51				
	Total	160	209	369	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	43	45	88	26	21	23	1.78	No	0.07	1.30
	Not Sure	25	31	56	15	14	15				
	Yes	97	141	238	59	65	62				
	Total	165	217	382	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	31	41	72	20	21	21	0.16	No	0.02	0.93
	Not Sure	11	13	24	7	7	7				
	Yes	116	138	254	73	72	73				
	Total	158	192	350	100	100	100				

Table B3.17. Experiments 1 and 2 Combined — Analysis of Alerts With and Without Text-to-Speech

Experiment	**1 and 2 Pooled**	**Tested Factor: Enhancement Feature - Text-to-Speech**										
H_0: Measured level of outcome construct is independent of whether an alert is delivered with Text-to-Speech or not. H_1: Delivery of alerts with Text-to-Speech improves the measured level of outcome construct												
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	29	16	45	8	4	6	7.94	**Yes**	0.10	1.79
	Partially	31	24	55	8	6	7				
	Yes	314	375	689	84	90	87				
	Total	374	415	789	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	106	105	211	30	26	28	1.99	No	0.05	1.21
	Not Sure	56	63	119	16	15	16				
	Yes	193	242	435	54	59	57				
	Total	355	410	765	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	232	281	513	67	70	68	4.69	No	0.08	0.68
	Somewhat	39	55	94	11	14	13				
	Yes	77	65	142	22	16	19				
	Total	348	401	749	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	118	103	221	34	26	30	6.63	**Yes**	0.09	0.96
	Not Sure	57	80	137	16	20	18				
	Yes	171	219	390	49	54	52				
	Total	346	402	748	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	92	97	189	30	24	26	4.60	No	0.08	1.15
	Not Sure	43	53	96	14	13	13				
	Yes	172	262	434	56	64	60				
	Total	307	412	719	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	70	102	172	24	27	25	0.76	No	0.03	0.92
	Not Sure	31	37	68	11	10	10				
	Yes	194	245	439	66	64	65				
	Total	295	384	679	100	100	100				

Table B3.18. Average Response Delay (Seconds) for Alerts With and Without Text-to-Speech

	Response Time Limit			
Feature	< 60 secs	< 100 secs	< 300 secs	No Limit
No TTS	34	42	65	10,514
With TTS	41	53	81	6,544



	Receipents who gave feedback	Receipents who viewed alert
Chi-Sq	350.43	528.37
C.V	0.38	0.46
O.R.	1.85	1.98

Figure 6.4. Response Rate for Alerts With and Without Text-to-Speech

B3.6. Location History

Table B3.19. Experiment 2 — Analysis of Geo-Targeted Alerts Filtered With and Without Location History

Experiment 2	**Tested Factor: Enhancement Feature - Location History**										
H_0: Measured level of outcome construct is independent of whether the alert was sent to users who frequently visit the geo-targeted area or not. H_1: Alerts that are targeted to recipients who frequently visit the geo-targeted alert area improve the measured level of outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency **0**	**1**	**Tot.**	Percentage **0**	**1**	**Tot.**	Chi-Sq.	Rej. H0?	Effect Size C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	7	5	12	3	4	3	0.45	NO	0.03	0.78
	Partially	6	3	9	2	2	2				
	Yes	256	123	379	95	94	95				
	Total	269	131	400	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	92	16	108	34	12	27	24.57	**YES**	0.25	3.35
	Not Sure	24	8	32	9	6	8				
	Yes	153	106	259	57	82	65				
	Total	269	130	399	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	178	92	270	66	70	68	1.25	NO	0.06	1.00
	Somewhat	37	13	50	14	10	13				
	Yes	53	26	79	20	20	20				
	Total	268	131	399	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	98	25	123	37	19	31	14.28	**YES**	0.19	2.19
	Not Sure	39	17	56	15	13	14				
	Yes	131	88	219	49	68	55				
	Total	268	130	398	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	89	30	119	33	23	30	5.19	NO	0.11	1.64
	Not Sure	34	14	48	13	11	12				
	Yes	147	86	233	54	66	58				
	Total	270	130	400	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	32	31	63	12	24	16	9.28	**YES**	0.15	0.56
	Not Sure	20	8	28	8	6	7				
	Yes	209	88	297	80	7	77				
	Total	261	127	388	100	100	100				
H_0: Subjects who are not moving towards the specific area targeted by an alert are equally interested, disinterested, and indifferent in receiving those alerts. H_1: More subjects who are not moving towards the specific area targeted by an alert are disinterested in receiving those alerts than those who are moving towards it.											
Hindsight Relevance: *In a real emergency, would you have preferred to receive this alert?*	**No**	55			20		20	147.09	**YES**	0.61 (Phi)	
	Not Sure	32			12		12				
	Yes	183			68		68				
	Total	270			100		100				

Table B3.20. Overall Impressions Regarding Geo-Targeted Alerts Filtered With and Without Location History

Feature Preference - Location History: Overall Impression (Usefulness)						
H_0: Subjects are equally divided in their overall impression regarding the usefulness of alerts which were filtered based on location history. H_1: More subjects find alerts filtered based on location prediction useful than those who are unsure or do not find them useful.						
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. H_0?	Effect Size Phi
Usefulness: *Some alerts targeted to a specific geophraphic area were delivered to you if you have recently visited the targeted area even if you were outside it when the alert was sent? Was this feature useful?*	**No**	3	6.4	44.09	**YES**	0.97
	Unsure	7	14.9			
	Yes	37	78.7			
	Total	47	100.0			

B3.7. Location Prediction

Table B3.21. Experiment 2 — Analysis of Geo-Targeted Alerts Filtered With and Without Location Prediction

Experiment 2	**Tested Factor: Enhancement Feature - Location Prediction**										
H_o: Measured level of outcome construct is independent of whether the alert was sent to users who are moving toward the alert area or not. H_1: Alerts that are targeted to recipients who are moving towards an alert area improve the measured level of outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding:	**No**	3		3	4	0	3	3.50	NO	0.17	1.97
Did you understand this alert message?	**Partially**		1	1	0	2	1				
	Yes	70	46	116	96	98	97				
	Total	73	47	120	100	100	100				
Relevance:	**No**	26	13	39	35	27	32	2.63	NO	0.15	1.75
In a real emergency, would this alert be relevant to you given your current situation and location?	**Not Sure**	5	1	6	7	2	5				
	Yes	43	34	77	58	71	63				
	Total	74	48	122	100	100	100				
Annoyance:	**No**	48	31	79	65	65	65	0.00	NO	0.00	1.00
In a real emergency, would this alert annoy you?	**Somewhat**	6	4	10	8	8	8				
	Yes	20	13	33	27	27	27				
	Total	74	48	122	100	100	100				
Actionability:	**No**	26	17	43	35	35	35	2.67	NO	0.15	1.34
In a real emergency, would this alert prompt you to take action?	**Not Sure**	7	1	8	9	2	7				
	Yes	41	30	71	55	63	58				
	Total	74	48	122	100	100	100				
Milling Behavior:	**No**	31	13	44	42	27	36	2.82	NO	0.15	1. 77
In a real emergency, would this alert prompt you to seek further information?	**Not Sure**	2	2	4	3	4	3				
	Yes	41	33	74	55	69	61				
	Total	74	48	122	100	100	100				
Adequacy:	**No**	7	2	9	10	4	8	2.99	NO	0.16	0.96
Did this alert contain enough information?	**Not Sure**	2	4	6	3	9	5				
	Yes	64	41	105	88	87	88				
	Total	73	47	120	100	100	100				
H_o: Subjects who are not moving towards the specific area targeted by an alert are equally interested, disinterested, and indifferent in receiving those alerts. H_1: More subjects who are not moving towards the specific area targeted by an alert are disinterested in receiving those alerts than those who are moving towards it.											
Hindsight Relevance:	**No**	23			31		31	44.95	**YES**	0.61 (Phi)	
In a real emergency, would you have preferred to receive this alert?	**Not Sure**	2			3		3				
	Yes	49			66		66				
	Total	74			100		100				

Table B3.22. Overall Impressions Regarding Geo-Targeted Alerts Filtered With and Without Location Prediction

<table>
<tr><th colspan="7">Feature Preference - Location Prediction: Overall Impression (Usefulness)</th></tr>
<tr><td colspan="7">H_o: Subjects are equally divided in their overall impression regarding the usefulness of alerts which were filtered based on location prediction.
H_1: More subjects find alerts filtered based on location prediction useful than those who are unsure or do not find them useful.</td></tr>
<tr><td rowspan="2">Outcome Construct:

Feedback Question</td><td rowspan="2">Answer (Level)</td><td rowspan="2">Frequency</td><td rowspan="2">Percentage</td><td rowspan="2">Chi-Sq.</td><td rowspan="2">Rej. H_0?</td><td>Effect Size</td></tr>
<tr><td>Phi</td></tr>
<tr><td rowspan="4">Usefulness:

Some alerts targeted to a specific geophraphic area were delivered to you if you were moving towards the targeted area even if you were outside it when the alert was sent? Was this feature useful?</td><td>No</td><td>4</td><td>8.5</td><td rowspan="4">39.9</td><td rowspan="4">YES</td><td rowspan="4">0.92</td></tr>
<tr><td>Unsure</td><td>7</td><td>14.9</td></tr>
<tr><td>Yes</td><td>36</td><td>76.6</td></tr>
<tr><td>Total</td><td>47</td><td>100.0</td></tr>
</table>

Appendix B4. Public Usability Trials (PUT) Data Tables and Statistical Analysis — Alert Characteristics

B4.1. Alert Timing

Table B4.1. Experiment 1 — Analysis of Alert Timing (Daytime vs. Nighttime Alerts)

Experiment 1	**Tested Factor: Alert Characteristic - Timing**										
H_o: Measured level of outcome construct is independent of whether an alert is issued during the day or night. H_1: Subjects respond to night-time and day-time alerts differently with respect to measured outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	3	2	5	4	3	4	0.8	No	0.08	1.76
	Partially	3	1	4	4	2	3				
	Yes	66	58	124	92	95	93				
	Total	72	61	133	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	17	17	34	24	28	26	0.3	No	0.04	0.88
	Unsure	11	9	20	15	15	15				
	Yes	43	35	78	61	57	59				
	Total	71	61	132	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	52	51	103	73	85	79	2.7	No	0.14	0.50
	Somewhat	8	4	12	11	7	9				
	Yes	11	5	16	15	8	12				
	Total	71	60	131	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	13	13	33	28	21	25	0.7	No	0.07	1.27
	Unsure	10	9	19	14	15	14				
	Yes	42	39	81	58	64	61				
	Total	72	61	133	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	19	15	34	27	25	26	2.8	No	0.15	0.71
	Unsure	6	11	17	8	18	13				
	Yes	46	34	80	65	57	61				
	Total	71	60	131	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	32	16	48	45	27	37	10.1	**Yes**	0.28	1.11
	Unsure	3	12	15	4	20	11				
	Yes	36	32	68	51	53	52				
	Total	71	60	131	100	100	100				

Table B4.2. Experiment 2 — Analysis of Alert Timing (Daytime vs. Nighttime Alerts)

Experiment 2	**Tested Factor: Alert Characteristic - Timing**											
H_0: Measured level of outcome construct is independent of whether an alert is issued during the day or night. H_1: Subjects respond to night-time and day-time alerts differently with respect to measured outcome construct.												
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	2	2	4	4	4	4	3.73	No	0.19	2.90
	Partially	6	1	7	11	2	7				
	Yes	45	49	94	85	94	90				
	Total	53	52	105	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	12	18	30	23	35	29	1.86	No	0.13	0.65
	Unsure	8	7	15	15	13	14				
	Yes	33	27	60	62	52	57				
	Total	53	52	105	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	32	24	56	60	47	54	2.74	No	0.16	1.05
	Somewhat	8	14	22	15	27	21				
	Yes	13	13	26	25	25	25				
	Total	53	51	104	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	13	18	31	25	35	30	3.12	No	0.17	0.50
	Unsure	9	12	21	17	24	20				
	Yes	31	21	52	58	41	50				
	Total	53	51	104	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	8	12	20	15	23	19	1.09	No	0.10	0.75
	Unsure	7	6	13	13	12	12				
	Yes	38	34	72	72	65	69				
	Total	53	52	105	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	17	13	30	35	27	31	0.83	No	0.09	1.30
	Unsure	3	4	7	6	8	7				
	Yes	28	31	59	58	65	61				
	Total	48	48	96	100	100	100				

Table B4.3. Experiments1 and 2 Combined — Analysis of Alert Timing (Daytime vs. Nighttime Alerts)

Experiment 1 and 2 Pooled2	**Tested Factor: Alert Characteristic - Timing**										
H_o: Measured level of outcome construct is independent of whether an alert is issued during the day or night. H_1: Subjects respond to night-time and day-time alerts differently with respect to measured outcome construct.											
Outcome Construct:	Answer	Frequency			Percentage			Chi-	Rej.	Effect Size	
Feedback Question	(Level)	**0**	**1**	**Tot.**	**0**	**1**	**Tot.**	Sq.	H0?	C.V	O.R.
Understanding:	**No**	8	5	13	4	4	4	5.60	No	0.13	2.33
Did you understand this alert message?	**Partially**	15	3	18	8	2	5				
	Yes	165	134	299	88	94	91				
	Total	188	142	330	100	100	100				
Relevance:	**No**	41	43	84	22	30	26	2.89	No	0.09	0.78
In a real emergency, would this alert be relevant to you given your current situation and location?	**Unsure**	29	19	48	16	13	15				
	Yes	116	80	196	62	56	60				
	Total	186	142	328	100	100	100				
Annoyance:	**No**	127	92	219	69	66	68	0.40	No	0.04	1.16
In a real emergency, would this alert annoy you?	**Somewhat**	27	22	49	15	16	15				
	Yes	29	25	54	16	18	17				
	Total	183	139	322	100	100	100				
Actionability:	**No**	46	41	87	25	29	27	0.95	No	0.05	0.81
In a real emergency, would this alert prompt you to take action?	**Unsure**	32	26	58	17	19	18				
	Yes	105	73	178	57	52	55				
	Total	183	140	323	100	100	100				
Milling Behavior:	**No**	41	35	76	22	25	23	2.13	No	0.08	0.73
In a real emergency, would this alert prompt you to seek further information?	**Unsure**	21	22	43	11	16	13				
	Yes	125	84	209	67	60	64				
	Total	187	141	328	100	100	100				
Adequacy:	**No**	55	33	88	31	25	28	4.00	No	0.11	1.03
Did this alert contain enough information?	**Unsure**	11	16	27	6	12	9				
	Yes	110	84	194	63	63	63				
	Total	176	133	309	100	100	100				

B4.2. Update Alert

Table B4.4. Experiment 1 — Analysis of Update Alerts vs. Initial Alerts

Experiment	**1**	**Tested Factor: Alert Characteristic - Update Alert**									
H_o: Measured level of outcome construct is independent of whether an alert is the first for an emergency situation or it is an update to a previous alert. H_1: Subjects respond to initial and update alerts differently with respect to measured outcome construct.											
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency			Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**Tot.**	**0**	**1**	**Tot.**			C.V	O.R.
Understanding: *Did you understand this alert message?*	**No**	16	2	18	9	2	6	4.5	No	0.13	1.58
	Partially	13	8	21	7	8	7				
	Yes	158	86	244	84	90	86				
	Total	187	96	283	100	100	100				
Relevance: *In a real emergency, would this alert be relevant to you given your current situation and location?*	**No**	50	37	87	27	39	31	4.2	No	0.12	0.65
	Unsure	26	12	38	14	13	14				
	Yes	109	46	155	59	48	55				
	Total	185	95	280	100	100	100				
Annoyance: *In a real emergency, would this alert annoy you?*	**No**	132	75	207	73	79	75	1.4	No	0.07	0.72
	Somewhat	17	7	24	9	7	9				
	Yes	33	13	46	18	14	17				
	Total	182	95	277	100	100	100				
Actionability: *In a real emergency, would this alert prompt you to take action?*	**No**	49	42	91	27	45	33	10.4	**Yes**	0.19	0.47
	Unsure	26	14	40	14	15	14				
	Yes	109	38	147	59	40	53				
	Total	184	94	278	100	100	100				
Milling Behavior: *In a real emergency, would this alert prompt you to seek further information?*	**No**	48	40	88	26	51	33	16.9	**Yes**	0.25	0.50
	Unsure	25	4	29	14	5	11				
	Yes	112	34	146	61	44	56				
	Total	185	78	263	100	100	100				
Adequacy: *Did this alert contain enough information?*	**No**	69	15	84	38	19	33	9.0	**Yes**	0.19	1.89
	Unsure	21	12	33	12	15	13				
	Yes	90	51	141	50	65	55				
	Total	180	78	258	100	100	100				

Appendix B5. Public Usability Trials (PUT) Data Tables and Statistical Analysis — Situational Awareness

As with other tested factors, contingency tables show both absolute and relative frequencies (percentages) of all outcome levels. The outcome levels are Wrong, Partially Correct,and Correct grades. The columns titled "0" represent the Wireless Emergency Alerts (WEA) group, and the columns titled "1" represent the Situation Digest feature. The contingency tables are followed by the Chi-square statistic for independence test, hypothesis testing verdict, and theoretical and practical effect sizes measured by Cramer's V (denoted as C.V) and odds ratio (denoted as O.R.), respectively. The odds ratio is calculated by dichotomizing the outcome levels as Correct grades (corresponding to a positive response) vs. Wrong grades (corresponding to a negative response). Partially Correct grades were disregarded in the calculation of the odds ratio.

B5.1. Type of Emergency

Table B5.1. Experiment 3 — Analysis of Situation Digest View vs. Normal WEA View With Respect to Correctly Identifying the Type of Emergency

<table>
<tr><td colspan="2">Experiment 3</td><td colspan="9">Tested Factor: Enhancement Feature - Situation Digest View</td></tr>
<tr><td colspan="2">Outcome Construct:
Situational Awareness - Type of Emergency</td><td colspan="9">Poll Question (Mid-Scenario):
An emergency is underway. What type of emergency is ii?</td></tr>
<tr><td colspan="11">H_0: Situational awareness as measured by level of outcome construct is independent of the view used to present alerts.
H_1: Subjects using Situation Digest view have improved situational awareness compared to regular WEA view.</td></tr>
<tr><td rowspan="2">Scenario</td><td rowspan="2">Answer (Level)</td><td colspan="2">Frequency</td><td colspan="3">Percentage</td><td rowspan="2">Chi-Sq.</td><td rowspan="2">Rej. H0?</td><td colspan="2">Effect Size</td></tr>
<tr><td>0</td><td>1</td><td>0</td><td>1</td><td>Tot.</td><td>C.V</td><td>O.R.</td></tr>
<tr><td rowspan="4">Earthquake with Plume</td><td>Wrong</td><td>50</td><td>20</td><td>50</td><td>20</td><td>35</td><td rowspan="4">18.87</td><td rowspan="4">Yes</td><td rowspan="4">0.31</td><td rowspan="4">3.87</td></tr>
<tr><td>P. Correct</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>Correct</td><td>51</td><td>79</td><td>50</td><td>80</td><td>65</td></tr>
<tr><td>Total</td><td>101</td><td>99</td><td>100</td><td>100</td><td>100</td></tr>
<tr><td rowspan="4">Random Alerts</td><td>Wrong</td><td>25</td><td>11</td><td>37</td><td>17</td><td>18</td><td rowspan="4">7.18</td><td rowspan="4">Yes</td><td rowspan="4">0.23</td><td rowspan="4">2.98</td></tr>
<tr><td>P. Correct</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>Correct</td><td>42</td><td>55</td><td>63</td><td>83</td><td>49</td></tr>
<tr><td>Total</td><td>67</td><td>66</td><td>100</td><td>100</td><td>100</td></tr>
<tr><td rowspan="4">Severe Weather</td><td>Wrong</td><td>46</td><td>29</td><td>69</td><td>44</td><td>38</td><td rowspan="4">3.30</td><td rowspan="4">No</td><td rowspan="4">0.11</td><td rowspan="4">1.59</td></tr>
<tr><td>P. Correct</td><td>73</td><td>77</td><td>109</td><td>117</td><td>75</td></tr>
<tr><td>Correct</td><td>14</td><td>14</td><td>21</td><td>21</td><td>14</td></tr>
<tr><td>Total</td><td>133</td><td>120</td><td>100</td><td>100</td><td>100</td></tr>
<tr><td rowspan="4">Alien Catastrophe</td><td>Wrong</td><td>118</td><td>78</td><td>49</td><td>33</td><td>41</td><td rowspan="4">34.18</td><td rowspan="4">Yes</td><td rowspan="4">0.27</td><td rowspan="4">27.23</td></tr>
<tr><td>P. Correct</td><td>101</td><td>160</td><td>50</td><td>60</td><td>55</td></tr>
<tr><td>Correct</td><td>1</td><td>18</td><td>1</td><td>7</td><td>4</td></tr>
<tr><td>Total</td><td>223</td><td>251</td><td>100</td><td>100</td><td>100</td></tr>
<tr><td rowspan="4">Bad Weather</td><td>Wrong</td><td>6</td><td>4</td><td>10</td><td>9</td><td>10</td><td rowspan="4">0.36</td><td rowspan="4">No</td><td rowspan="4">0.06</td><td rowspan="4">1.50</td></tr>
<tr><td>P. Correct</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>Correct</td><td>48</td><td>48</td><td>90</td><td>91</td><td>90</td></tr>
<tr><td>Total</td><td>54</td><td>52</td><td>100</td><td>100</td><td>100</td></tr>
<tr><td rowspan="4">All Scenarios</td><td>Wrong</td><td>245</td><td>142</td><td>43</td><td>24</td><td>33</td><td rowspan="4">45.90</td><td rowspan="4">Yes</td><td rowspan="4">0.20</td><td rowspan="4">2.37</td></tr>
<tr><td>P. Correct</td><td>174</td><td>237</td><td>30</td><td>40</td><td>35</td></tr>
<tr><td>Correct</td><td>156</td><td>214</td><td>27</td><td>36</td><td>32</td></tr>
<tr><td>Total</td><td>575</td><td>593</td><td>100</td><td>100</td><td>100</td></tr>
</table>

B5.2. Action to Perform

Table B5.2. Experiment 3 — Analysis of Situation Digest View vs. Normal WEA View With Respect to Correctly Identifying the Action to Perform

Experiment 3		**Tested Factor: Enhancement Feature - Situation Digest View**								
Outcome Construct: **Situational Awareness - Action**		Poll Question (Mid-Scenario): *An emergency is underway. What should you do now?*								
H_0: Situational awareness as measured by level of outcome construct is independent of the view used to present alerts. H_1: Subjects using Situation Digest view have improved situational awareness compared to regular WEA view.										
Scenario	Answer (Level)	Frequency		Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**0**	**1**	**Tot.**			C.V	O.R.
Earthquake with Plume	**Wrong**	45	26	45	26	36	7.31	**Yes**	0.19	2.26
	P. Correct	0	0	0	0	0				
	Correct	56	73	55	74	65				
	Total	101	99	100	100	100				
Random Alerts	**Wrong**	41	18	61	27	44	15.50	**Yes**	0.34	4.21
	P. Correct	0	0	0	0	0				
	Correct	26	48	39	73	56				
	Total	67	66	100	100	100				
Severe Weather	**Wrong**	66	27	49	22	36	21.69	**Yes**	0.29	4.73
	P. Correct	54	67	40	54	47				
	Correct	15	29	11	24	17				
	Total	135	123	100	100	100				
Alien Catastrophe	**Wrong**	141	127	63	51	57	11.91	**Yes**	0.16	5.55
	P. Correct	79	109	35	43	40				
	Correct	3	15	1	6	4				
	Total	223	251	100	100	100				
Bad Weather	**Wrong**	30	34	10	12	11	0.00	No	0.00	1.01
	P. Correct	0	0	0	0	0				
	Correct	265	257	90	88	89				
	Total	295	291	100	100	100				
All Scenarios	**Wrong**	316	220	55	37	46	35.64	**Yes**	0.17	2.14
	P. Correct	133	176	23	30	26				
	Correct	131	195	23	33	28				
	Total	580	591	100	100	100				

B5.3. Immediacy of Emergency

Table B5.3. Experiment 3 — Analysis of Situation Digest View vs. Normal WEA View With Respect to Correctly Identifying the Immediacy of Emergency

Experiment 3		**Tested Factor: EnhancemenFeature - Situation Digest View**								
Outcome Construct: **Situational Awareness - Immediacy**		Poll Question (Mid-Scenario): *An emergency is underway. When do you need to take action?*								
H_0: Situational awareness as measured by level of outcome construct is independent of the view used to present alerts. H_1: Subjects using Situation Digest view have improved situational awareness compared to regular WEA view.										
Scenario	Answer (Level)	Frequency		Percentage			Chi-Sq.	Rej. H0?	Effect Size	
		0	**1**	**0**	**1**	**Tot.**			C.V	O.R.
Earthquake with Plume	**Wrong**	70	41	69	41	56	15.75	**Yes**	0.28	3.19
	P. Correct	0	0	0	0	0				
	Correct	31	58	31	59	45				
	Total	101	99	100	100	100				
Random Alerts	**Wrong**	37	35	55	53	54	0.06	No	0.02	1.09
	P. Correct	0	0	0	0	0				
	Correct	30	31	45	47	46				
	Total	67	66	100	100	100				
Severe Weather	**Wrong**	89	70	47	40	44	2.32	No	0.08	1.27
	P. Correct	19	24	10	14	12				
	Correct	81	81	43	46	45				
	Total	189	175	100	100	100				
Alien Catastrophe	**Wrong**	132	154	45	45	45	3.86	No	0.08	1.29
	P. Correct	100	98	45	29	31				
	Correct	59	89	20	26	23				
	Total	291	341	100	100	100				
Bad Weather	**Wrong**	30	34	10	12	11	0.35	No	0.02	0.86
	P. Correct	0	0	0	0	0				
	Correct	265	257	90	88	89				
	Total	295	291	100	100	100				
All Scenarios	**Wrong**	358	334	38	34	36	2.98	No	0.04	1.19
	P. Correct	119	122	13	13	13				
	Correct	466	516	49	53	51				
	Total	943	972	100	100	100				

B5.4. User Preference for Situation Digest View

Table B5.14. Overall Perceptions Regarding the Situation Digest View

<table>
<tr><th colspan="7">Feature Preference — Situation Digest View: Overall Impression</th></tr>
<tr><td colspan="7">H_0: Subjects do not care about how alert messages from multiple ongoing interrelated incidents are presented to them on their phones.
H_1: More subjects prefer the Situation Digest view than those who prefer the Normal WEA view (List of Messages), are indifferent or are unsure.</td></tr>
<tr><td rowspan="2">Outcome Construct:

Feedback Question</td><td rowspan="2">Answer (Level)</td><td rowspan="2">Frequency</td><td rowspan="2">Percentage</td><td rowspan="2">Chi-Sq.</td><td rowspan="2">Rej. H_0?</td><td>Effect Size</td></tr>
<tr><td>Phi</td></tr>
<tr><td rowspan="5">Preference:

You were presented with more than one view to represent the messages. One was a stream of messages and the other was an updating situation panel. Which one do you prefer?</td><td>Situation Digest View</td><td>18</td><td>49</td><td rowspan="5">11.5</td><td rowspan="5">Yes</td><td rowspan="5">0.56</td></tr>
<tr><td>Normal View (List of Messages)</td><td>5</td><td>14</td></tr>
<tr><td>Unsure</td><td>6</td><td>16</td></tr>
<tr><td>Both</td><td>8</td><td>22</td></tr>
<tr><td>Total</td><td>37</td><td>100</td></tr>
</table>

APPENDIX B6. PUBLIC USABILITY TRIALS (PUT) DATA TABLES AND STATISTICAL ANALYSIS — OVERALL IMPRESSIONS ABOUT WEA BENEFITS

Table B6.1. Experiment 1 Results for Overall Impressions Regarding WEA Benefits

Experiment 1	**Overall Perception - WEA Benefits**					
H_0: After frequent exposure to WEA alerts, subjects are equally divided among their beliefs regarding usefulness and lifesaving potential of WEA. H_1: More subjects believe that WEA is useful and can save lives than those who believe that it is not useful, cannot save lives, or are unsure about its benefits.						
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. Ho?	Effect Size Phi
Usefulness: *Do you believe wireless emergency alerts are useful?*	**Unsure**	0	0.0	38.0	**YES**	1.41
	Yes	19	100.0			
	No	0	0.0			
	Total	19	100.0			
Usefulness: *Do you believe wireless emergency alerts could save lives?*	**Unsure**	0	0.0	40.0	**YES**	1.41
	Yes	20	100.0			
	No	0	0.0			
	Total	20	100.0			

Table B6.2. Experiment 2 Results for Overall Impressions Regarding WEA Benefits

Experiment 2	**Overall Perception - WEA Benefits**					
H_0: After frequent exposure to WEA alerts, subjects are equally divided among their beliefs regarding usefulness and lifesaving potential of WEA. H_1: More subjects believe that WEA is useful and can save lives than those who believe that it is not useful, cannot save lives, or are unsure about its benefits.						
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. Ho?	Effect Size Phi
Usefulness: *Do you believe wireless emergency alerts are useful?*	**Unsure**	6	12.8	26.1	**YES**	0.74
	Yes	41	87.2			
	No	0	0.0			
	Total	47	100.0			
Usefulness: *Do you believe wireless emergency alerts could save lives?*	**Unsure**	4	8.5	32.4	**YES**	0.83
	Yes	43	91.5			
	No	0	0.0			
	Total	47	100.0			

Table B6.3. Experiment 3 Results for Overall Impressions Regarding WEA Benefits

Experiment 3	**Overall Perception - WEA Benefits**					
H_0: After frequent exposure to WEA alerts, subjects are equally divided among their beliefs regarding usefulness and lifesaving potential of WEA. H_1: More subjects believe that WEA is useful and can save lives than those who believe that it is not useful, cannot save lives, or are unsure about its benefits.						
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. Ho?	Effect Size Phi
Usefulness: *Do you believe wireless emergency alerts are useful?*	**Unsure**	0	0.0	74.0	**YES**	1.41
	Yes	37	100.0			
	No	0	0.0			
	Total	37	100.0			
Usefulness: *Do you believe wireless emergency alerts could save lives?*	**Unsure**	0	0.0	74.0	**YES**	1.41
	Yes	37	100.0			
	No	0	0.0			
	Total	37	100.0			

Table B6.4. Pooled Results for Overall Impressions Regarding WEA Benefits

Experiment 1,2,3 Pooled	**Overall Perception - WEA Benefits**					
H_0: After frequent exposure to WEA alerts, subjects are equally divided among their beliefs regarding usefulness and lifesaving potential of WEA. H_1: More subjects believe that WEA is useful and can save lives than those who believe that it is not useful, cannot save lives, or are unsure about its benefits.						
Outcome Construct: *Feedback Question*	Answer (Level)	Frequency	Percentage	Chi-Sq.	Rej. H_0?	Effect Size Phi
Usefulness: *Do you believe wireless emergency alerts are useful?*	**Unsure**	6	5.8	172.1	**YES**	1.29
	Yes	97	94.2			
	No	0	0.0			
	Total	103	100.0			
Usefulness: *Do you believe wireless emergency alerts could save lives?*	**Unsure**	4	3.8	184.9	**YES**	1.33
	Yes	100	96.2			
	No	0	0.0			
	Total	104	100.0			