**ATIS/SIP Forum IP NNI Task Force**

**Tyson's Corner, VA**

**November 7-8, 2017**

**Contribution**

**TITLE: Usability Test Results**

**SOURCE\*: Jonathan Nelson (Hiya)**

**ISSUE NUMBER:**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Abstract**

This contribution provides study descriptions, results, and conclusions on the testing of warning phrasing and iconography. Display recommendations are also made based on the results.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**ATIS-0x0000x**

ATIS Standard on

**ATIS Technical Report on a Framework for Display of Verified Caller ID**

**Alliance for Telecommunications Industry Solutions**

Approved Month DD, YYYY

**Foreword**

The Alliance for Telecommunications Industry Solutions (ATIS) serves the public through improved understanding between carriers, customers, and manufacturers. The [**COMMITTEE NAME**] Committee [**INSERT MISSION**]. [**INSERT SCOPE**].

The mandatory requirements are designated by the word *shall* and recommendations by the word *should*. Where both a mandatory requirement and a recommendation are specified for the same criterion, the recommendation represents a goal currently identifiable as having distinct compatibility or performance advantages. The word *may* denotes a optional capability that could augment the standard. The standard is fully functional without the incorporation of this optional capability.

Suggestions for improvement of this document are welcome. They should be sent to the Alliance for Telecommunications Industry Solutions, [**COMMITTEE NAME**], 1200 G Street NW, Suite 500, Washington, DC 20005.

At the time of consensus on this document, [**COMMITTEE NAME**], which was responsible for its development, had the following leadership:

[**LEADERSHIP LIST**]

The **[SUBCOMMITTEE NAME]** Subcommittee was responsible for the development of this document.

**Revision History**

| **Date** | **Version** | **Description** | **Author** |
| --- | --- | --- | --- |
|  |  |  |  |

**Table of Contents**

[INSERT]

**Table of Figures**

[INSERT]

**Table of Tables**

[INSERT]

# Scope, Purpose, & Application

## Scope

## Purpose

## Application

.

# Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this Standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this Standard are encouraged to investigate the possibility of applying the most recent editions of the standards indicated below.

# Definitions, Acronyms, & Abbreviations

For a list of common communications terms and definitions, please visit the *ATIS Telecom Glossary*, which is located at < <http://www.atis.org/glossary> >.

## Definitions

## Acronyms & Abbreviations

# Architecture

Editor’s note: add figure illustrating various access technologies and a variety of device types (UEs).

# Signaling of Verified Caller ID using Conventional Caller Name (CNAM)

## Considerations during Transition to STIR/SHAKEN Caller Authentication and Signaling

# Display Guidelines of Verified Caller ID on All-IP Networks and Screen-based Devices

## Entities that shape the display

## Assumptions

## Available Call-related Information

## Preliminary Display Usability and Comprehension Studies by Hiya

Over the course of Q2 and Q3 2017, Hiya Inc. conducted several usability studies targeted at the display guidelines of the STIR/SHAKEN protocol. The goals of the study were twofold:

1. To measure the potential impact of a positive assurance indication for verified calls (the “green checkmark”), and
2. To assess various textual and iconographic display options for caution indicators on suspicious calls.

Hiya conducted three independent brief usability studies on different audiences:

1. A user comprehension and influence study on robust caller profiles and certified call markers;
2. A user impact analysis of various phrasings and iconography for suspicious call messaging; and.
3. A call pickup rate impact analysis of a “certified” checkmark icon against existing Hiya users.

### Study #1: Comprehension and Impact of Certified Call Markers

#### Study description:

Hiya interviewed 11 randomly chosen individuals among a pool of volunteers to participate in a paper-based usability study.

Participants were offered no context or explanation of STIR/SHAKEN, or the purpose of the study.

Participants were presented with a sequence of incoming call UI mockups (see Exhibit A), and asked a series of questions:

* What do you notice about this screen?
* What confidence or trust do you have in this caller?
* What effect does this have on your opinion of previous screens (if any)?

The mockups were designed to enhance the incoming call experience with expanded caller profile information, with the final mockup showing a “certified call” confirmation reflective of STIR/SHAKEN validation.

Interviews lasted an average of 35 minutes per participant.

#### Results:

In nearly all cases (10/11), participants showed a “strong” or “very strong” indication that additional caller profile information strengthened their confidence in the legitimacy of the theoretical phone call.

This applied across all mockups, prior to “certified call” marker.

For the “certified call” mark,

* A significant number of participants (8/11) identified the mark as reassurance to the legitimacy of the call, and all who identified this also expressed future doubts to the legitimacy of any future call lacking the indication.
* Some (6/11) expressly stated that they would become less likely to answer any unknown call lacking certification, once the certification marker was seen.

#### Conclusions:

Hiya has concluded that the use of such a marker with clear context would negatively impact all calls lacking this marker.

It remained unclear from the study if the marker would have a positive impact on pickup rates for marked numbers. (Hiya has previously demonstrated that extended call profile information positively influences call pickup rates, without any “certified call” indication.)

### Study #2: Warning Phrasing Test

#### Study Description

Hiya has crafted 7 near-identical mockups of an incoming call screen. The only delta between the screen was the presence of a warning phrase, expressing reasonable doubt about the legitimacy of the caller. Study participants were shown this screen for 5 seconds, then asked a series of questions:

1. Would you answer this call?
2. Was this call from a trustworthy source?
3. Would you block this number from calling in the future?
4. Why do you think you’re receiving this call?

The following phrases were used, each shown to 400 unique participants (no overlap between phrases):

* Phone number only (baseline)
* “Fake Phone Number” with phone number
* “Possible Fraud” with phone number
* “Private Number” with no phone number
* “Unknown Caller” with phone number
* “Caller Not Verified” with phone number
* “Spoofed Number” with phone number

#### Results

Results were analyzed as a delta from the baseline (number-only) test, with these results:

|  |  |  |  |
| --- | --- | --- | --- |
| **Display** | **Pickup Rate** | **Block Rate** | **User Trust** |
| Phone number only | 30% | 29% | 46% |

Individually, the results compared to the baseline were:

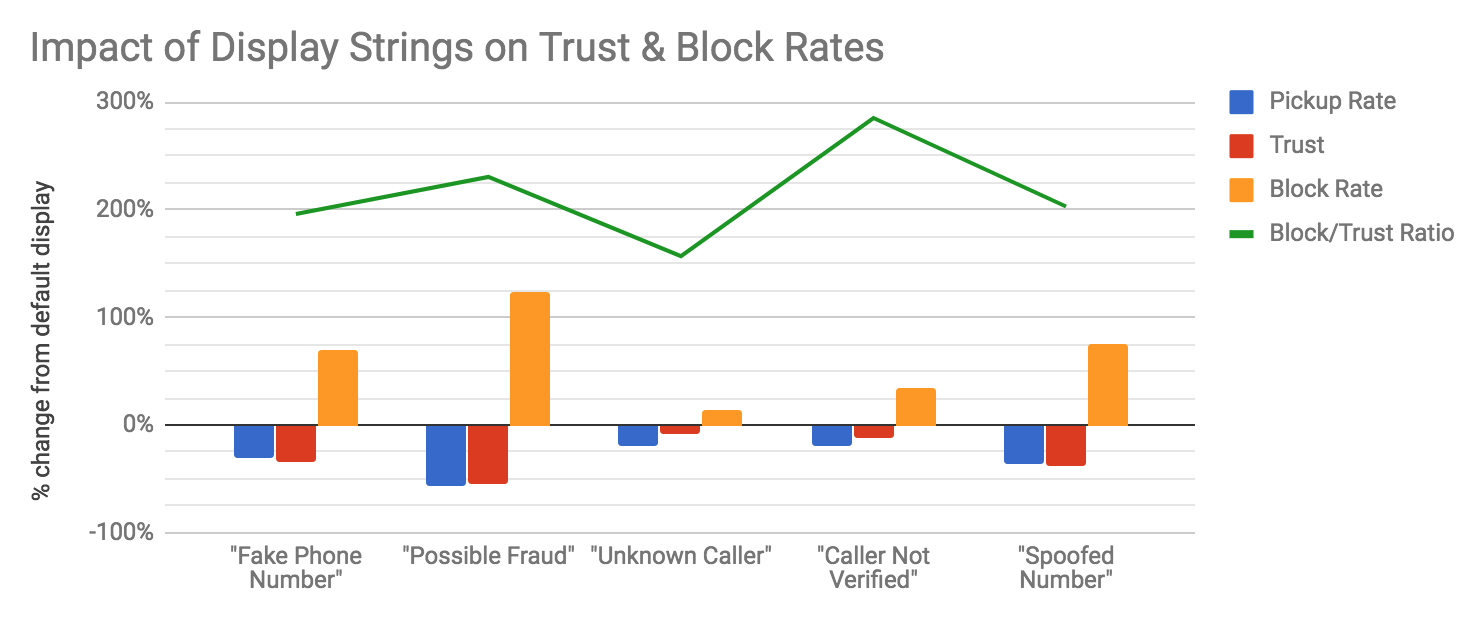
|  |  |  |  |
| --- | --- | --- | --- |
| **Display** | **Pickup Rate** | **Block Rate** | **User Trust** |
| “Private Number”  (no phone number) | -7% | +48% | -16% |
| “Fake Phone Number” | -30% | +69% | -35% |
| “Possible Fraud” | -57% | +124% | -54% |
| “Unknown Caller” | -20% | +14% | -9% |
| “Caller Not Verified” | -20% | +34% | -12% |
| “Spoofed Number” | -37% | +76% | -37% |

#### Conclusions

One main caveat is that these numbers should be validated against actual user behavior. A 29% block rate for unknown callers seems unexpectedly high. This should be verified via other means before concluding this reflects real-world behavior. However, while totals may not be actionable, the belief is the deltas between levels are valid.

“Private Number” was included in this study only as an anecdotal comparison and is not a recommended display string. From this we can see fairly minimal change in pickup rates, but a large increase in block rates. It was otherwise excluded from analysis.

The main goal of the study is to find how different messaging affects users’ level of caution for an incoming call, while separately measuring their likelihood to block the number in the future. The ideal message would greatly increase caution level, with minimal impact on block rates (as spoofed calls may impersonate legitimate numbers).



*Figure: Decrease in user trust, increase in block percentages, and ratio between the two*

Overall, we observed some variation in the delta of user trust as it relates to the delta in block rates. Less aggressive message “Unknown Caller” has a lower block/trust ratio, but overall has negligible impact on both. Most influential warning “Possible Fraud” has over 2:1 ratio. Finding an ideal string may require additional tests, and individual service providers might determine appropriateness based on prominence/availability of call blocking capabilities.

Similar results of “Fake Phone Number” and “Spoofed Number” indicates user general understanding of the term “spoof”.

### Study #3: Call Pickup Rate Test

#### Study Description

Hiya selected 70 high-volume (in excess of 600 monthly observed calls) legitimate business phone numbers from Canada with varying pre-existing caller profile information. Canadian numbers were chosen because of Hiya’s strong adoption rate of the Samsung “Smart Call” caller ID service.

Call creation rates and user pickup rates were monitored for these numbers over a two-week period, during which Hiya caused the display of a simple white checkmark to be shown for any Smart Call user.

After two weeks, the checkmark was removed and call volume and pickup rates were measured for an additional two weeks.

#### Results

It is important to note that Hiya could not identify or control the individuals receiving phone calls from the selected businesses. Therefore, it is not possible to confirm the results of the initial study (that individual pickup rates would drop as users become familiar with expecting a checkmark on legitimate calls).

On average, Hiya observed a 4.89% increase in call pickup rates overall while the checkmark was present.

#### Conclusions

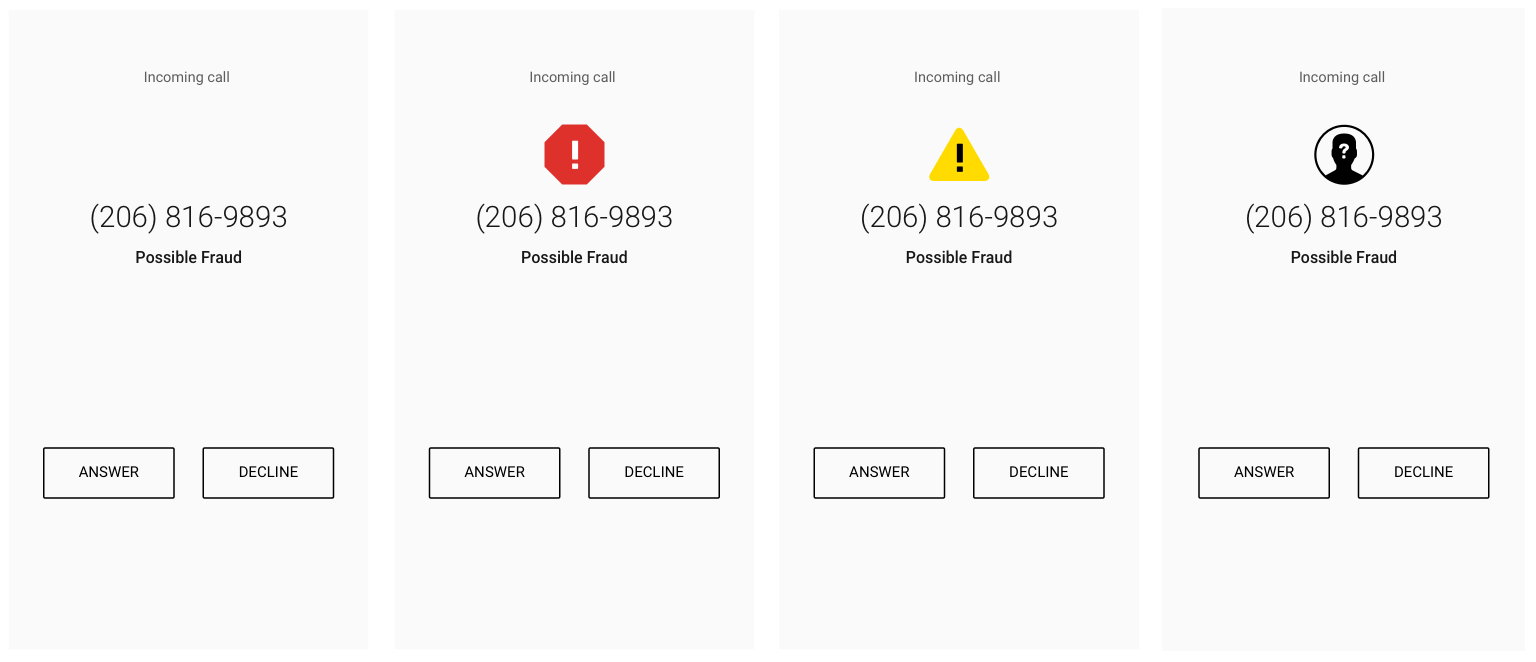
Further analysis is pending to determine if the number’s existing call profile (a “caller ID” name) or business industry was more effective than others.

But overall, this is taken to reinforce the positive impact a certified marker can potentially have on pickup rates for verified calls.

### Study #4: Iconography Impact Test

#### Study Description

Hiya selected two test strings “Possible Fraud” and “Fake Number”, and created 8 mockups of an incoming call screen, 4 with each string. Each string was paired with one of four possible flagged icons: no icon, a red stop sign, a yellow triangle, and an “unknown person” icon.



*Figure: Test images for “Possible Fraud” string*

Each participant (400 per mockup with no overlap) was asked the same questions as Study #2:

1. Would you answer this call?
2. Was this call from a trustworthy source?
3. Would you block this number from calling in the future?
4. Why do you think you’re receiving this call?

#### Results

Results were analyzed as a delta from the no-icon entities, with these results:

|  |  |  |  |
| --- | --- | --- | --- |
| **Display** | **Pickup Rate** | **Block Rate** | **User Trust** |
| “Possible Fraud” | 13% | 65% | 21% |
| “Fake Number” | 21% | 49% | 30% |

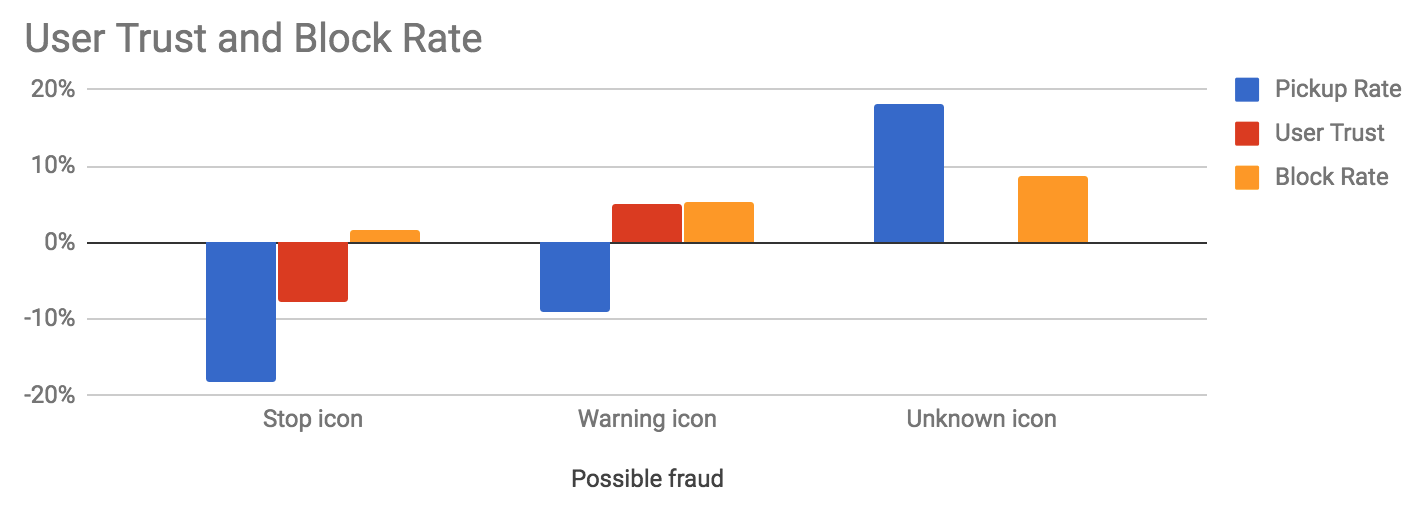
With the addition of an icon, the following deltas from baseline were observed:

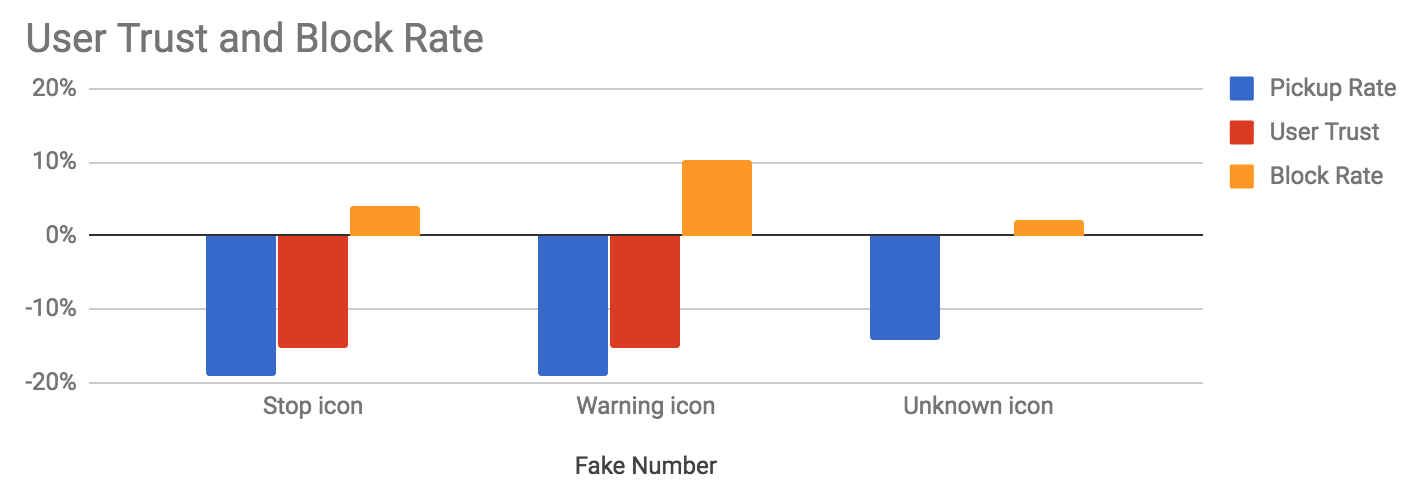
|  |  |  |  |
| --- | --- | --- | --- |
| **Display** | **Pickup Rate** | **Block Rate** | **User Trust** |
| “Possible Fraud” + stop sign  “Fake Number” + stop sign | -18%  -19% | +2%  +4% | -8%  -15% |
| “Possible Fraud” + warning sign  “Fake Number” + warning sign | -9%  -19% | +5%  +10% | +5%  -15% |
| “Possible Fraud” + unknown sign  “Fake Number” + unknown sign | +18%  -14% | +9%  +2% | 0%  0% |

#### Conclusions

Further analysis is recommended based on these results. Initial observations:

* The addition of icons has relatively minor impact on block rates, above that from text alone. In all cases though, block rates increased further.
* With the less severe “fake number” string, a supporting icon more strongly reduces pickup rates and overall perception of trustworthiness of the caller.
* With the more severe “possible fraud” string, the less-severe “unknown” icon actually improved pickup rates, tempering the impact of the warning message.





*Figure: Impact on user trust and block rates, by icon.*

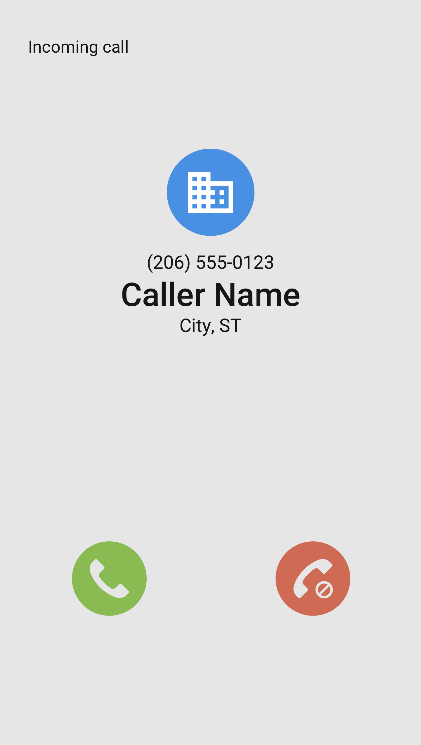
## Recommended Data Treatment and Display Options

## Example Displays

These examples are provided for the illustration and enhancement of the scenarios listed in the above table.

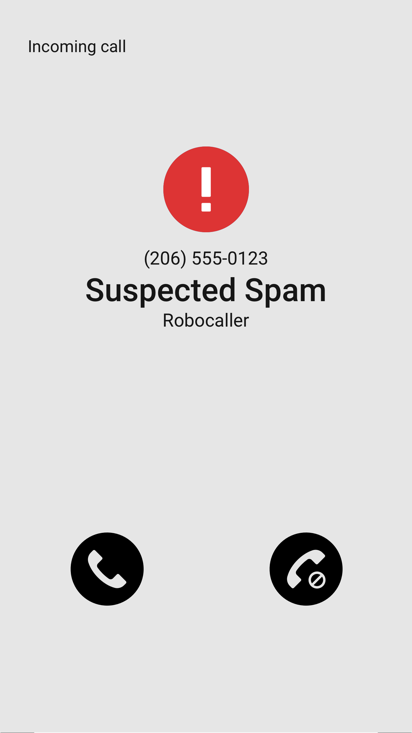
### Full Attestation and Verification Passed (no analytics)

In this scenario, the user does not subscribe to a CVT service. The delivery of the “verstat” parameter (TN validation passed) delivers the call to the UE without warnings (or affirmations). The logo and location (city and state) of the caller is retrieved and delivered by enhanced CNAM (eCNAM).



### Gateway Attestation, Verification Passed, subscribes to analytics (analytics determine the call is suspicious)

In this scenario, the user subscribes to a CVT service that provides analytics. A gateway attestation is inconclusive to the caller, but an analytics service has flagged the caller as a known scammer. Therefore, a warning is provided to the user.

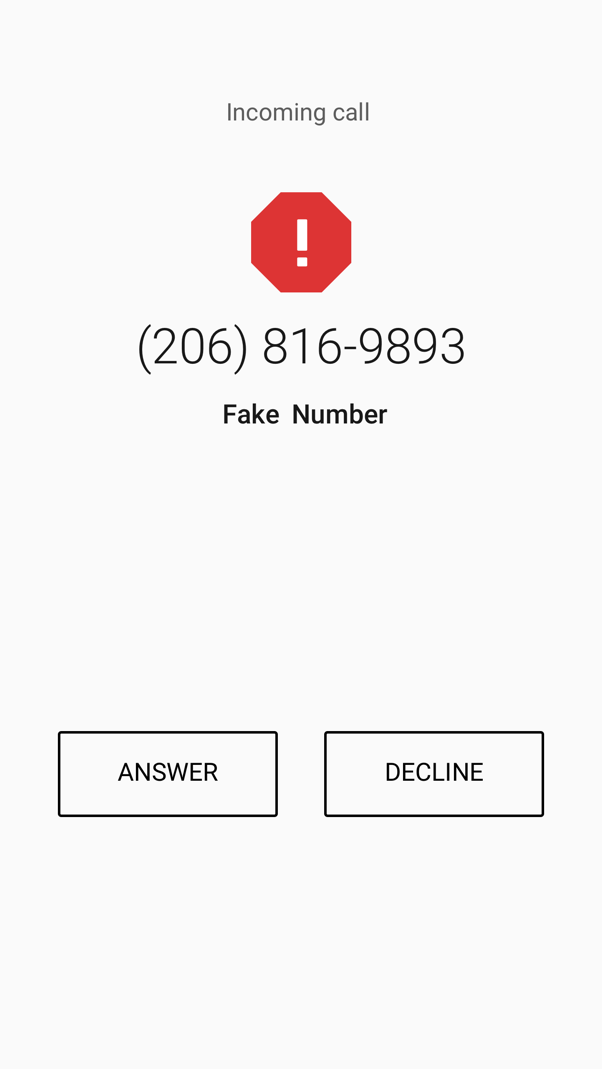


### Verification Failed

If the verification failed, CVT may not be necessary or be used. There are several possible outcomes.

The service provider may simply block the call from terminating to the end user, per the end user's request.

Alternatively, a warning would be provided along with an explanation of the reason behind the warning. The end user is then forewarned and empowered to manage incoming calls based on all the information made available. A CVT service may be able to provide more useful reasoning, but is not expected to.



### Basic recommendations on the Display or Message delivery to the UE

As a result of the above studies, it is recommended to adopt the following display/messaging behavior:

1. On validation failure, recommend use of words “Fake Number” to inform user not to trust the number being presented. If an icon can be presented, use of a “stop sign” further increases user caution with negligible impact on block rates. String is short enough to fit in CNAM fields.
2. For UEs that include blocking capability, recommend cautioning user that blocking this number is not blocking the actual caller, or disallow blocking in cases of known spoofing.
3. eCNAM delivers the aggregate of all the information available about the TN (caller identity, results of CVT analytics, and information queried by the terminating provider)
4. The use of multiple symbols in a given display is not recommended because the consumer's interpretation of different symbols may result in confusion and detract from the value the service is providing.
5. Displaying status symbols, such as checkmarks, on calls with "full attestation – verification passed", is not recommended (studies show it leads to consumer confusion).
6. It is recommended that only warning symbols be provided when warranted.
7. Audible special ringing/tones may be applied on calls that fail verification as a consumer option.

## ADA Considerations

# Display Guidelines for Analog Devices

## Analog Devices connected to an IP Network

# Related SDOs and Fora

## 3GPP

## GSMA

## Cable Labs

## Consumer Electronics

# Conclusions

**Annex A**

(normative/informative)

# A Illustrative Examples

This annex will document supportive material



