情報処理学会論文誌 Vol.61 No.3 726–736 (Mar. 2020)

推薦論文

「かわいい」画像を用いた行動誘引による セキュリティ警告の効果改善

皆川 諒[1,†1,a)] 高田 哲司[1,b)]

受付日 2019年5月15日, 採録日 2019年11月29日

概要：セキュリティ警告はセキュリティ脅威への遭遇可能性をICT利用者に伝え，それを回避可能にするための仕組みである．しかし，利用者の多くはセキュリティ警告を有効活用していないという現状がある．そこで本研究では，セキュリティ警告の効果改善を目的とし，「かわいい」と感じる視覚的刺激による行動誘引効果をセキュリティ警告に応用することを提案する．このアイデアに基づき，セキュリティ警告内に「かわいい効果」をもたらしうる画像を導入したプロトタイプ警告を実装し，それを用いて実験参加者による評価実験を行った．その結果，かわいい効果を導入した警告は既存の “そうではない” 警告と比較してセキュリティ警告内の警告文を理解するよう利用者を誘引できる，という結果を得た．

キーワード：ユーザブル・セキュリティ，セキュリティ警告，かわいい，行動誘引，ユーザ・インタフェース

Improving the Effect of Security Warning with a Stimulus of “Kawaii”

Ryo Minakawa[1,†1,a)] Tetsuji Takada[1,b)]

Received: May 15, 2019, Accepted: November 29, 2019

Abstract: Security warning dialog is a security function to inform ICT users about the potential risk of encountering a security threat and makes it possible to avoid it. However, most users do not make use of the dialog as the dialog provider intended. We propose to apply the action attraction effect caused by the visual stimulus that feels “cute” to the security warning dialog for the effective use of the dialog. Based on the idea, we implemented a prototype warning dialog with an image that could bring a cute effect, and conducted an evaluation experiment using the dialog with subjects from crowdsourcing. From the experiment, we obtained the result that the warning dialog with a cute image can significantly attract the user to understand the warning message in the dialog compared with the dialog without a cute image.

Keywords: usable security, security warning, security alert, kawaii, user interface, psychology

1. はじめに

「セキュリティ警告」(Security warning，Security alert) と呼ばれる仕組みが存在する．これは，ICTシステムの利用時にセキュリティ脅威・侵害に遭遇する可能性を利用者に通知し，適切な対応を促すための仕組みである．著名なものとしては，不適切な電子証明書によるTLS通信に関わる警告や，ダウンロードしたファイルを実行する際に表示される警告などがあげられる．

これらのセキュリティ警告を不要にすることは現時点では困難だと考える．理想としては，なんらかのセキュリティシステムが “脅威判定” と “なすべき対応” を自動で確実に実行してくれることが望ましい．しかし，以下にあげる理由からそれは困難である．

[1] 電気通信大学
The University of Electro-Communications, Chofu, Tokyo 182–8585, Japan
[†1] 現在，NTT Communications Corporation
Presently with NTT Communications Corporation
[a)] r.mina1124@gmail.com
[b)] zetaka@computer.org

本論文の内容は2018年10月のコンピュータセキュリティシンポジウム2018にて報告され，同プログラム委員長により情報処理学会論文誌ジャーナルへの掲載が推薦された論文である．

© 2020 Information Processing Society of Japan

- 誤検知・検知不能の問題：既存のセキュリティシステムには誤検知の問題が存在し，現時点で存在するセキュリティ上の脅威を100%正確に検知するのは困難である．また，新たな脅威が次々と発見されるため，その検知が不能な脅威が存在しうることがあげられる．
- なすべき対応の決定：利用者の利用用途や利用環境に応じてなすべき対応が1つに定まらない場合がある．よって，セキュリティシステムによる自動対応だと支障をきたす可能性がある．

したがって，利用者が置かれている状況に関する情報提供ととりうる対応の選択肢を提示し，最終的な判断を利用者に促すのが「セキュリティ警告」の役割である．

しかし，セキュリティ警告はその有効性が疑問視されている．Sunshineらが2008年に行ったWebブラウザのSSL警告に関するオンライン調査[8]によると，30〜60%のユーザが警告に遭遇しても，それを無視してWebサイトにアクセスするという結果を報告している．またKrolらの論文[14]でも，PDFファイルのダウンロード時における警告を用いて評価を行った結果，80%以上の実験参加者が警告を無視した，という結果を報告している．このような状況は，セキュリティ警告が設置者の意図したとおりに機能していないことを示しており，情報セキュリティ対策の仕組みとして望ましくない．

この問題に対して我々は，セキュリティ警告に「かわいい」効果を加えることを提案する[1]．具体的には，セキュリティ警告の画面に「かわいい」と利用者が思うであろう画像を一緒に提示する，というシンプルな方法である．この提案を行った理由は，セキュリティ警告に対して心理的な誘引効果を付与することにより，セキュリティ警告に対する注目効果を発揮し，かつ冷静な判断を促しうると考えたためである．ここでいう誘引効果とは，ある種の刺激により特定の行動・効果が誘発される，と定義する．「かわいい」による心理的効果については，Nittonoらの研究[19]を根拠にしている．この研究における実験では，タスクを行う前にかわいい画像を見せた場合と見せない場合とでタスクの実行結果に有意な差が生じた．この結果は「かわいい画像」を見たことによる心理的効果によるものであり，かわいい画像を見た実験参加者は，集中力が高まり，かつ注意深く振舞ったからであると結論付けている．また一般的な傾向として，人が“かわいい”と思うものを見ると「近づきたくなる，注視する，触りたくなる，大切に扱う」といった行動が誘引されることは多くの人が直感的に理解できることだろう．我々の提案はこれを論拠とし，セキュリティ警告への応用可能性について研究を行ったものである．

以降本論文では，2章でセキュリティ警告の問題改善に取り組んだ先行研究について述べ，3章では我々が提案する“かわいい効果”付きセキュリティ警告について紹介する．4章では，提案手法の効果検証を目的とした評価実験の手法と結果について述べ，5章では，実験結果に関する考察と今後の課題について述べる．

2. 関連研究

セキュリティ警告を対象とする先行研究について述べる．

警告内警告文の改善：論文[8], [9]は，どちらもWebブラウザにおけるSSL警告に関する研究である．Sunshineらの研究[8]は，既存のWebブラウザに実装されているSSL警告を取り上げ，その効果について比較評価を行っている．Feltらの研究[9]では，SSL警告における情報提示法について，どのように文字ベースの情報提供を行うべきかについて改良を試み，その検証を行っている．

利用者への提示刺激追加：セキュリティ警告は，一般にダイアログ形式で情報提供は文字により行われていた．これに対し，文字や色，アイコン画像による装飾を行うことで，利用者の注意をひきつけるよう提示刺激を増やすという工夫は行われてきた．

Andersonらの研究[17]では，これらのほかに，警告ダイアログに回転，揺れそしてWindowサイズの拡大縮小といったアニメーション効果を付与することで提示刺激の追加を試み，それがセキュリティ警告に対する馴化の抑制に効果があることを示した．またBalebakoらの研究[20]では，音をセキュリティ警告に応用し，警告の認知率改善と視覚による警告の見落とし防止に効果があることを示した．

警告への対応方法の変更：セキュリティ警告に対する対応は，その警告内にあるボタンをクリックするのが一般的である．これはセキュリティ警告がダイアログという形態で実装されているためである．この状況に対し，この対応方法を工夫することによりセキュリティ警告の効果改善を試みた研究が存在する．

Bravo-Lilloらの研究[11], [12]では，セキュリティ警告への対話手法を工夫することによりセキュリティ警告の問題改善を試みた．提案されている対話手法としては，1) 警告が出現してから一定時間が経過しないと警告の操作ができない方法，2) ボタンクリックを複数回行わせる方法，3) 警告ダイアログ内の重要情報をスワイプさせる方法，4) 警告ダイアログ内の重要情報をキー入力させる方法，である．これらの方法は，一定の有効性があることを評価を通じて示しているが，操作時間や対話操作によるコストが欠点となることも指摘している．

藤原らの研究[18]では，不快感を利用した警告インタフェースを提案している．その実装として，キーボードとマウスなど複数の操作を必要とする操作法や，ボタンクリックを通常とは異なる“マウス右ボタン”で行わせる，選択した処理が実際に実行されるまで待ち時間を課す，といった対話法を提案している．

警告の出現タイミングの改良：セキュリティ警告は，警告出現前のタスクを中断させる形で利用者に対応を要求す

© 2020 Information Processing Society of Japan

る．それがセキュリティ警告における問題を引き起こしているとし，その改善を試みる研究が行われている．

Jenkins らの研究 [16] では，警告ダイアログを出現させるタイミングを制御し，警告が出現する前のタスクをなるべく中断させないタイミングで警告を提示することにより，警告を無視する事態を減少させる可能性があることを明らかにした．

一方，Weinberger らの研究 [15] では，HTTPS のエラー警告を見てそれを無視すると利用者が決めた場合，その警告を再び出現させるまでの期間はどうあるべきか？という問題を提起し，その問題に対して 6 つのポリシを提案するとともに，実験参加者により評価検証を行った研究である．

外部の評価情報の追加：Yang らの研究 [13] では，フィッシング Web ページの警告に対し，その Web ページに関する “traffic ranking” を評価情報として付与することにより，なぜその警告が表示されたのかを利用者に理解できるようにするという工夫を提案している．

警告の設計ガイドライン：Bauer らの研究グループは，これまでの研究成果をふまえ，警告ダイアログの画面設計に関する 6 つのガイドラインを示した [5]．

3. 提案するセキュリティ警告

本章では，我々が提案する新たなセキュリティ警告の改善手法について述べ，かつその方針に基づき実装したプロトタイプシステムを紹介する．

3.1 設計方針

本研究では，前章で述べた先行研究のうち「利用者への提示刺激追加」のアプローチについて新たな手法を提案する．文字情報による情報提示法の改善や，警告内情報の増大に対してどこが重要な情報かを示し，その認識を確実にする視覚的工夫など，様々な試みが行われていることは前章で述べたとおりである．しかし，これらのアプローチは利用者に対して様々な刺激を追加することにより「なんとかして対応行動を実行させる」というアプローチであるといえる．このアプローチだと，どれだけの刺激を追加すればどれだけの利用者が対応行動を行うかを見極める必要が生じる．また馴化の問題により，利用者がその刺激に順応して対応行動をとらなくなる懸念も残る．そのため，新たな刺激を次々と追加する必要が生じるかもしれない．

そこで本研究では，問題改善のために特定の行動を引き出すような刺激を警告に適用する．つまり，多種または多様な刺激を提示することにより利用者が行動するよう仕向けるのではなく，ある刺激により思わずある行動を行ってしまう，という効果をセキュリティ警告に応用するのである．ここで “ある刺激” として我々は “かわいい” という感情を発生させる画像をセキュリティ警告に適用する．1 章でも述べたが，人が “かわいい” と思うものを見ると「注視する，近づきたくなる，触りたくなる，大切に扱う」といった行動が誘発されることは多くの人が理解できると考える．こういった行動がセキュリティ警告出現時に発現すれば，警告への対応行動が改善される可能性があると考える．また，かわいい画像が人の振舞いにどのような影響を及ぼすのかを検証した研究として Nittono らの研究 [19] がある．この研究では 3 つの評価実験から，人がかわいい画像を見ることでポジティブな感情が発生し，それが注意深く振る舞うようにさせると結論付けている．本研究ではこれらを論拠とし，その効果のセキュリティ警告への応用可能性を検証する．

またもう 1 つの設計方針として，セキュリティ警告利用時の利便性を可能な限り損なわないよう配慮することとした．先行研究におけるアプローチの 1 つに「警告への対応方法の変更」がある．これらは警告への対応行動に対して一定の時間経過を強制したり，ボタンを複数回押下，キー入力，タッチパネルの操作やそれらの操作の組み合わせで対応行動を行わせるようにする．このアプローチは警告への理解を促進することが確認されているが，一方で操作時間や操作手間をコストとして利用者に負担させることになる．つまり，これらのアプローチで得られた警告効果の改善は操作コストを利用者に課すことにより得られたものであり，操作コストと警告効果のトレードオフとなっている．これをふまえ，我々は新たな操作コストを発生させないまま，上記提案により警告効果が改善されるほうが望ましいと考え，この方針を決定した．

3.2 プロトタイプシステム

前述の方針に基づき実装したセキュリティ警告のプロトタイプシステムについて述べる．前述の設計方針に基づき実装した「かわいい」効果付きセキュリティ警告を**図 1** に示す．矩形状ウィンドウ内にダイアログ形式のユーザインタフェースとして実装している点は既存のセキュリティ警告と同様である．既存の警告との差異は，かわいい効果を発現させるため警告ダイアログ内左上に「子犬の画像」を提示している点のみである．このように既存の警告と提案するプロトタイプ警告との差異を「かわいい効果」に寄与する部分だけにすることで，既存警告との効果を比較検証可能にする．

なおプロトタイプ警告のインタフェースで「かわいい効果」以外の部分については Bauer らが提唱するガイドライン [5] における設計指針に基づいている．推奨された対応行動を行うボタンを，そうでないボタンよりも大きくする，警告内の文章は簡潔にしつつも提示情報の詳細を知りたいユーザに対しては詳細情報を取得可能にする，といった画面設計はそれにしたがったものである．

© 2020 Information Processing Society of Japan

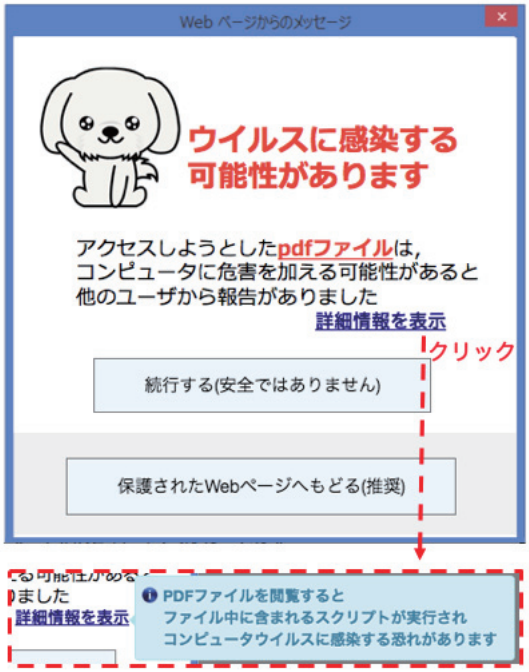

図 1 かわいい効果付きセキュリティ警告
Fig. 1 Kawaii effect applied in security warning dialog.

4. 評価実験

提案手法の効果を検証するため，クラウドソーシングを通じて実験参加者を募集し，オンライン実験による評価を実施した．本章では実施した評価実験について，実験方法とその結果を述べる．

4.1 実験方法

本節では実験方法について説明する．本実験は，実験参加者によるオンライン実験として実施した．クラウドソーシングを通じて実験参加者を募り，実験実施側が指示するタスクを Web ブラウザを通じて実験参加者に行わせる方法である．利用したクラウドソーシングは Yahoo クラウドソーシング [3] である．タスク内容は，「情報セキュリティに関するオンラインアンケート（設問数 10）」とし，タスク完了に対する報酬は，T-Point 10 ポイントとした．なお実験参加の条件は「Google Chrome ブラウザを利用し，オンラインで実験タスクを実施できる環境を持つこと」とし，それ以外の条件は設けなかった．なお，このタスク内容が本実験における目的と異なる理由については後述する．

この依頼タスクに応募してきた実験参加者に対し，実験実施側は比較考察のため 3 種の警告プロトタイプを用意し，その 1 つをランダムに割り当てた．割り当ては実験システムを通じて自動的に行った．したがって，本実験は “between-subject design” による実験であり，各実験参加者は 3 種の警告のうちの 1 つにのみ遭遇したことになる．なお 3 種の警告プロトタイプシステムについては 4.2 節で述べる．

実験手順の概要は以下のとおりである．

(1) **事前説明**：実験内容について，Web ページと書類を閲覧していただく形で実験参加者に説明を行った．なお前述のとおり，実験内容の説明は本来の実験目的ではなく，みかけ上の実験目的として「情報セキュリティに関する意識と実際に行っている対策の実態に関するアンケート調査」と説明した．これはセキュリティ警告に対する利用者の反応を測定するため，それを事前に説明してしまうと利用者の対応行動に影響をおよぼす恐れがあったためである．

(2) **タスク実施**：実験実施側が用意したオンラインアンケート（設問数 10）」を依頼した．これは，情報セキュリティを専門とする大学研究室がクラウドソーシングを通じて依頼するタスクとして違和感のないように見せるためである．なお，このアンケート回答の途中，9 問目の回答のあとに実験参加者に割り当てたセキュリティ警告を 1 回出現させた．この警告に対する実験参加者の反応について，実験システムを通じて測定を行った．

(3) **実験に関する事後アンケート**：依頼タスクとしてのアンケートへの回答終了後，今回の実験ならびに実験参加者の属性に関して事後アンケート調査を実施した．

(4) **事後説明**：実験の真の目的が「セキュリティ警告に対する利用者の反応に関する評価」であったこと，また実験結果に影響をおよぼす恐れがあったため，それを事前に説明せずに実験を実施したことを説明し，あらためて実験参加への同意を取得した．

実験における測定値と測定方法，ならびに事後アンケートの内容については，4.4 節の実験結果とともに説明する．

4.2 実験に用いたセキュリティ警告プロトタイプ

評価実験に用いたセキュリティ警告のプロトタイプシステムについて説明する．今回の実験では，比較検証のため以下の 3 種のセキュリティ警告を実装して実験を行った（**図 2**）．

条件 1）「かわいい画像」付きのセキュリティ警告
条件 2）「かわいさのない画像」付きのセキュリティ警告
条件 3）既存のセキュリティ警告

3 つの警告とも基本的な外観は同じであり，差分は警告内左上に表示される画像のみとなっている．なお条件 3 は，ベースラインとなる警告を想定したものであり，条件 1 で “かわいい効果” による警告への反応を測定することを意図しつつ，条件 2 で “かわいさのない画像” でも “かわいい画像” と同等の効果が得られるかを検証することを意図した条件設定となっている．

© 2020 Information Processing Society of Japan

図 2 評価実験に用いた 3 種の警告画面：差分は左上の画像のみである
Fig. 2 Three types of security warnings for the experiment.

表 1 実験参加者の人数と性別比，専門知識を有する参加者数
Table 1 Participant information.

	実験完了者数	男女比（男性：女性）	専門知識あり
条件 1	29 名	22:7	9 名（31.0%）
条件 2	22 名	13:9	5 名（22.7%）
条件 3	33 名	24:9	8 名（24.2%）

表 2 実験参加者の年齢分布
Table 2 Age distribution of the participants.

20 代	30 代	40 代	50 代	60 代以上
4 名	20 名	37 名	20 名	3 名

4.3 実験参加者

本実験では各警告条件に 50 名の参加者を割り当てる予定で 150 名の実験参加者を募集した．しかし，4.1 節の手順どおりに実験を行い，かつ検証に必要な測定データが得られた参加者数は合計 84 名となった（**表 1**）．また表 1 内には，各条件に割り当てられた参加者群の男女比，ならびに IT 専門知識を持つとアンケートで回答した参加者の人数も示している．なお本実験では，IT 専門知識の有無を「プログラミング言語の学習経験」とし事後アンケートの回答によって判定した．また**表 2** に実験参加者の年齢分布を示す．

応募してきた実験参加者のうち，測定データが完全に揃わない参加者が 64 名になった理由は，以下の 2 つである．

- 実験条件に記載された Web ブラウザを利用してタスクを実施しなかった．
- 参加条件を満たさない場合，タスクが実施できない実装ではなかった．

前述のとおり，実験参加者募集の際に実験条件として Google Chrome ブラウザの使用を明記していたが，それをきちんと理解せずにそれとは異なる Web ブラウザでタスクを行ってしまった人が多数いたと推測している．よって以降の議論では，この 84 名による評価結果を有効回答として議論を進める．

表 3 警告に対する反応
Table 3 User response to security warning dialog.

	推奨行動選択率（人数）	警告文再生率（人数）
条件 1	93.1%（27 名）	51.7%（15 名）
条件 2	90.9%（20 名）	40.9%（ 9 名）
条件 3	79.8%（26 名）	21.2%（ 7 名）

表 4 警告への反応に要した時間
Table 4 The response time to the three warning types.

	median（s）	mean（s）	S.D.（s）
条件 1	14.74	15.88	6.38
条件 2	15.86	22.26	17.16
条件 3	15.55	19.65	12.67

4.4 実験結果

本節では，警告に対する反応として測定された結果と，実験終了後に行った事後アンケートの結果について述べる．

4.4.1 警告に対する反応

警告に対する利用者の反応として，以下の値を実験システムを通じて測定した．

- 対応行動：推奨行動を選択したかどうか？（2 値）
- 対応時間：警告が表示されてから，警告内のボタンを押下するまでの時間（数値）

実験結果を**表 3**，**表 4** に示す．表 3 は，対応行動に関する結果について推奨行動をとった実験参加者の割合と人数を各条件ごとに示している．表 4 は，対応時間に関する結果について 3 条件の警告に対する実験参加者の反応時間を中央値，平均値，標準偏差で示している．また**図 3** は，各条件における対応時間のデータ分布を箱ひげ図として示している．

4.4.2 事後アンケートとその結果

依頼タスク完了後に行った事後アンケート内容とその結果について述べる．

(1) 警告への対応行動とその自己認識：実験参加者が警告に対して実際に行った対応行動について，警告から情報を得て，自分で判断し，対応行動を選択したかを検証する

© 2020 Information Processing Society of Japan

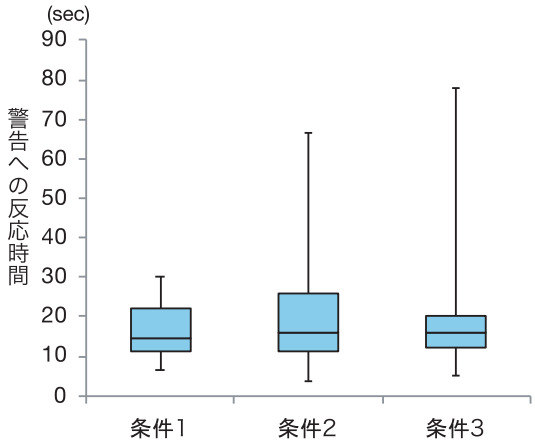

図 3 警告への対応時間の分布
Fig. 3 Distribution of response time in a box plot.

ため，警告に対して行った対応行動を事後アンケートで聞き直した．実際には推奨行動を選択していたにもかかわらず，事後アンケートで推奨行動をとらなかった，と回答したり，実際の対応にかかわらず，この設問にはどちらを選択したか覚えていない，と回答した場合は，警告に対して適切に対応したとはいいがたい，と判定する．アンケートは，(推奨行動に従わなかった（上ボタン)，推奨行動に従った（下ボタン)，覚えていない）の 3 択で行った．

結果としては，各条件に 1 名ずつ計 3 名の実験参加者が自分が行った実際の対応とアンケート回答が異なる状況に該当した．また彼らの警告に対する対応時間は第一四分位数よりも短い反応時間という点も共通していた．したがってこの 3 名は，警告をよく見ずに拙速に対応行動を選択してしまったものと考える．

なお 4.1 節の手順 (2) から手順 (3) までにかかる時間は，実験参加者が意図的に時間間隔をとらない限り，長くても 5 分前後と推測する．また警告との遭遇は 1 回だけであるため，「警告への対応内容を忘れた」という状況は起こりにくいと考えている．この結果から，今回の実験で得られたデータのほとんどは，提示された警告をきちんと認識したうえで対応行動を選択した結果であると考える．

(2) 警告内メッセージの理解：事後アンケートにおいて，実験で遭遇した警告の中に書かれていた「警告文」をきちんと認識しているかについて 6 択の選択肢から回答させた．6 つの選択肢のうち 4 つは 4 種の警告文例を提示し，残りの 2 つは「4 種の警告文のいずれも異なる」と「わからない，覚えていない」とした．正解は 4 種の警告文例の中に含まれており，回答結果は（1. 正解/2. 不正解）として判定し，集計した．

結果は表 3 の最右列に示している．警告文をきちんと認識していた実験参加者の割合は，21.2〜51.7%となった．この結果について，ベースライン条件である条件 3 と他の 2 条件についてフィッシャーの正確確率検定を行ったところ，p 値はそれぞれ，条件 1–3 間が $p = 0.0171$，条件 2–3 間が $p = 0.1390$ となった．したがって，条件 1–3 間の場合のみ “有意差あり”（$p = 0.017 < 0.05(*)$）という結果となった．

表 5 かわいい効果の検証結果
Table 5 User impression to the image in each warning.

	かわいい	どちらでもない	かわいくない	スコア平均
条件 1	19	8	2	0.586
条件 2	1	7	14	−0.591

表 6 提示した警告を偽警告と疑ったか？
Table 6 Did a user suspect the presented dialog as fake?.

	偽警告を知っていた	疑った	疑わなかった
条件 1	29 名	24 名	5 名
条件 2	22 名	14 名	8 名
条件 3	29 名	27 名	2 名

(3) 警告内画像の “かわいい効果”：警告内左上に表示した画像について，“かわいい効果” が意図どおりに発生したかについて事後アンケートを通じて検証した．検証方法は，(1. かわいい，2. かわいくない，3. どちらでもない）の 3 つの選択肢を提示し，条件 1 の警告に遭遇した実験参加者は，警告内画像を見て「かわいい」と思ったかを，条件 2 の警告の場合は，「かわいいとは思わなかった」かを検証した．

検証結果を**表 5** に示す．この結果から条件 1 の警告内画像に対しては，65.5%の参加者が警告内画像を “かわいい” と感じていた．また，回答結果を，(かわいい，どちらでもない，かわいくない）＝（1，0，−1）として数値化し，スコア化した回答の平均値を表内の最右列に示している．この値を用いて t 検定を実施したところ，p 値は，$p = 6.7 \times 10^{-9} \ll 0.01(**)$ となった．したがって，条件 1 と条件 2 の警告内画像は，“かわいい効果” について異なる効果を発生させていたといえる．

(4) 偽警告との疑い：近年，警告画面を模倣した画面を表示し，偽物のセキュリティ対策ソフトウェアのインストールを促したり，サポート窓口に電話させるなどの被害が発生している [10]．今回の実験でも，依頼タスクであるアンケートの回答途中に警告を表示したため，実験参加者はこれを「偽の警告」と判断し，その判断に基づいて対応行動を決定した可能性がある．その影響を測るため事後アンケートによる調査を行った．調査は，(1. 偽警告だと疑った，2. 偽警告だと疑わなかった，3. 偽警告の存在を知らなかった）の 3 択で行った．

結果を**表 6** に示す．実験参加者のうち「偽警告」を知らなかった人が 4 名いて，その全員が条件 3 の参加者であった．よって表 6 内の左から 2 列目「偽警告を知っていた」に書かれた人数が本分析の母集団となる．この状況において，偽警告と疑った参加者の割合は，(条件 1，条件 2，条

© 2020 Information Processing Society of Japan

表 7 出現した警告を実験の一部と認識したか？
Table 7 Did you recognize the appeared warnings as part of the experiment?.

	一部だと思った	一部だと思わなかった	どちらともいえない
条件 1	19（65.5%）	6	4
条件 2	19（86.4%）	1	2
条件 3	25（75.8%）	7	1
推奨行動群	56（76.7%）	10	7
非推奨行動群	7（63.6%）	4	0

件 3 = 82.8%，63.6%，93.1%）となり，全条件において 6 割を超える実験参加者が「実験中に提示された警告を偽警告と疑った」という結果となった．

(5) 出現した警告を実験の一部だととらえたか？：本実験では，実験参加者の反応に対する影響を回避するため，実験前に本来の実験目的を参加者に説明しなかった．しかし，実験を通じて実験参加者が説明された実験目的とは別の目的があるのではと考え，それを理由に警告への対応を決定した可能性がある．それを検証するため，事後アンケートで「出現したセキュリティ警告は実験の一部だと思ったか？」について調査した．回答は（1. 一部だと思った，2. 一部だと思わなかった，3. どちらともいえない）の 3 択で実施した．

結果を**表 7** に示す．この表では警告の条件別ならびに対応行動の推奨行動および非推奨行動を選択した参加者群という観点で結果を集計している．この結果から各条件における実験参加者の 6 割以上が出現した警告を「実験の一部」と認識して実験を行っていたことが分かった．なお表 7 内の 3 条件間ならびに対応行動 2 群間で，本来の実験目的の気づきについて差が生じたかをフィッシャーの正確確率検定で検証した．その結果，警告の 3 条件間は，$p = 0.2070$，対応行動の 2 群間は，$p = 0.1787$ となり，いずれの場合も “有意差なし” という結果となった．したがって，特定の警告により「別の研究目的の存在に気づかれた」ということはなかったと考える．また対応行動が別の研究目的の存在に気づいたことにより影響を受けているともいいがたい，と判断した．

(6) 実験参加者の衝動性：警告への反応について実験参加者の性格的な要因が影響を及ぼしている可能性があるとの報告がある．論文 [4] では，フィッシングサイトの回避能力と心理特性の関係性を報告しており，セキュリティ警告でも同様の関係性が存在する可能性があると考え，本調査を行った．小橋らの論文 [6] を参考に，被験者の衝動性をアンケートを通じて調査した．調査は，8 問の設問を 4 段階で回答させるものであり，我々はこの回答を単純な方法でスコア化して利用した．スコア（S）の値域は，（$8 \leq S \leq 32$）（最小値 $1 \times 8 = 8$，最大値 $4 \times 8 = 32$）であり，スコアが大きいほど「衝動性が高い」ということになる．

表 8 警告への対応行動と衝動性
Table 8 Response to warnings and impulsivity of participants.

	母集団人数	スコア平均	スコアの標準偏差
推奨行動群	73	16.10	3.46
非推奨行動群	11	18.18	3.87

表 9 専門知識有群による警告への対応行動
Table 9 Expertise of participants and response action.

	対象人数	推奨行動	非推奨行動
条件 1	9	8	1
条件 2	5	4	1
条件 3	8	8	0
合計	22	20	2

結果を**表 8** に示す．表内の 2 群について t 検定を行ったところ，$p = 0.0350 < 0.05(*)$ となり有意差ありという結果となった．したがって，衝動性が高い人が非推奨行動を取る傾向にある，ということが明らかになった．

(7) IT 専門知識の有無：IT 専門知識の有無が警告への反応に影響を及ぼす可能性を調査するため，実験参加者にプログラミング言語の学習経験を問うことで IT 専門知識の有無を調査した．各警告条件において IT 専門知識を有していると申告した実験参加者の人数は，表 1 の最右列に記載のとおりである．この専門知識を有していた実験参加者 22 名を母集団とし，各条件ごとに推奨行動群と非推奨行動群に分類した結果を**表 9** に示す

この結果から，専門知識を有している参加者が推奨行動をとった割合は 90.9%（= 20/22）となる．これに対して専門知識を持っていない参加者が推奨行動をとった割合は，85.5%（= 53/62）となる．このことから，警告への対応行動が今回の判定方法による IT 専門知識の有無によって影響を受けているとはいいがたい．

5. 考察

5.1 かわいい効果による警告効果の改善

今回の評価実験により，“かわいい効果” について以下の 2 点が明らかになった．

結果 1) 警告文の認識を確実にさせる効果がある．
結果 2) 警告への対応行動を改善する効果があるとはいえない．

結果 1) は，4.4.2 項「(2) 警告内メッセージの理解」の調査結果から導いたものである．条件 1 の「かわいい効果付き警告」は，他の 2 条件の警告よりも有意に警告文を再生させることができていたことが明らかになっており，このことから「かわいい効果付き警告」は利用者の注意を警告に引きつけ，警告文を確実に認識させる，という点について効果を発揮しうる可能性があることが明らかになったと考える．

結果 2) は，実験システムを通じて測定した 2 種の客観

© 2020 Information Processing Society of Japan

的データの分析結果から導かれたものである．分析結果を以下に示す．

- 対応行動における改善効果：
  フィッシャーの正確確率検定で，3 つの警告条件間に有意差なし（$p = 0.2315$）．
- 対応時間における改善効果：
  One-way ANOVA による検定で，3 つの警告条件間に有意差なし（$p = 0.2884$）．

これらの結果より，警告への対応行動を改善する効果があるとはいえない，という結論になった．ただし，この 2 つの分析において有意差が生じなかったことについては，以下に述べる 2 つの要因が影響していた可能性があると我々は考えている．

1 つめの要因は「“かわいい効果” により誘引される振る舞いは，警告への対応行動決定において決定的な影響を及ぼすほどではなかった」というものである．警告への対応の決定においては，今回導入した効果の影響だけでなく，それ以外の要因も影響するからである．代表的なものとしては，警告遭遇時の状況，ユーザの該当領域における知識，セキュリティ警告に対するユーザの経験，印象，信用などが考えられる（文献 [7] 参照）．よってセキュリティ警告への対応行動を決定する判断プロセスにおける各種要因の影響についてあらためて調査・考察し，そのうえで “かわいい効果” がそれに良い影響を発揮しうるかを検討しなおす必要があると考えている．これについては今後の課題とする．

2 つめの要因は，実験設計に起因するものである．今回の実験では，各警告条件で少なくとも約 80%の実験参加者が推奨行動を選択した，という結果になっている（表 3）．これは，警告条件によらず，実験参加者の多くが「推奨行動」を選択し，結果として「警告条件間で有意差なし」という結果になったものと考えられる．つまり，リスク回避と報酬獲得の二者択一の状況において，多くの実験参加者がリスク回避を選択するに至る「何か」が存在したと考えられ，それは実験設計にあると考えている．

その可能性の 1 つとして「非推奨行動を選択する理由が少なかった」という点が考えられる．リスク回避よりも報酬獲得を選択するだけの動機付けが弱く，結果的に警告条件によらず推奨行動が選択された結果であると推測する．この点については，タスク報酬額を変更する，非推奨行動を選択する傾向の強い状況下で評価するなど，実験設計の修正を試みる必要があると考えている．

もう 1 つの可能性としては，実験参加者の属性にかたよりがあったと考えられる．今回の実験参加者は，多くが情報セキュリティに対して安全志向の強い方々であったと推測される．それゆえ，警告条件によらず警告への対応行動は「推奨行動」を選択することになり，結果として有意差を示す結果には至らなかった，と考える．これに対する対策としては，実験参加者の募集時に一定の要件を設ける，事前に実験参加者の属性を調査し，その結果に応じて実験を割り振るなどの方法が考えられる．これも今後の課題とする．

なお，結果 2) において有意差がなかった原因は，“かわいい効果” そのものがなかったからではないか？ということも可能性として考えられる．これについて我々は，そのような事態であったとは考えにくいと認識している．その根拠は以下のとおりである．

根拠 a) 表 5 の結果より，警告内に提示された画像を “かわいい” と感じた実験参加者が存在すること

根拠 b) 結果 1) より，警告メッセージの認識において “かわいい効果” による改善効果がみられたこと

根拠 c) 文献 [19] より，著者ら以外にもその効果を示す結果が提示されていること

これらの点から，“かわいい効果” そのものがなかったと解釈するのは難しいと考えており，「一定の効果を実験参加者におよぼしていた」が「警告に対する対応行動を推奨行動に誘導するほどの影響力にはならなかった」と解釈するのが妥当であると考える．

5.2 実験設計と限界

本実験における実験方法は「理想とする実験環境」とは異なっていた．この点が実験結果に望ましくない形で影響をおよぼしていた可能性がある．理想の実験環境は「実際にセキュリティ脅威に直面する状況でセキュリティ警告に遭遇した」と実験参加者が状況認識し，その警告に対応することである．しかしながら，事後アンケートの (5)「実験目的への気づき」の結果から，実験参加者の 75%が実験中に出現した警告を「評価実験の一部」と認識していたことが明らかになった．さらに，非推奨行動を選択した参加者 11 名のうち 4 名は「実験だから安全が確保されていると考えた」と非推奨行動を選択した理由を述べている．また事後アンケートの (4)「偽警告との疑い」でも，各警告条件で 60%以上の参加者が出現した警告を偽警告だと疑ったと回答している．これらの結果は，今回の実験が理想の実験環境とは異なる状況であったことを示しており，今回の実験結果はこの点をふまえる必要がある．

しかし，これらの要因を低減・排除する実験設計には限界があると考えている．「実験目的への気づき」について，参加者がそれに気づくことを排除することは困難であろう．事前の実験説明を真の目的とは異なる内容で行っても，実験中にその目的とは関連の薄い事象が発生することで違和感を覚え，「実験の真の目的は説明内容とは別にあるのでは？」と実験参加者が考えることを排除する実験設計は容易ではないと著者らは考えている．

出現したセキュリティ警告を「偽警告と疑う」点についても，同様であると考えている．被験者の知識や経験に依存する部分はあるものの，評価対象の警告画面を「100%偽

© 2020 Information Processing Society of Japan

警告ではない」と確信を持たせたうえで評価を行うには「セキュリティ警告に対する評価を行う」と事前説明をする以外の方法では困難だと考えている．しかし，それを行うことは評価実験として望ましくない影響が発生しうることは前述したとおりである．

これらのことからいえるのは，セキュリティ警告の評価方法はどのようにしたら妥当といえるのかを検討していく必要があるという点である．現実におけるセキュリティ警告との遭遇場面と同等の状況をどの程度まで実験設計で作り込むべきなのか？は今後の課題の 1 つであると著者らは考えている．

5.3 今後の課題

以下の 4 つを今後の課題とし，今後も評価実験を継続する予定である．

(1) 実験設計の修正と実験参加者の振り分け：本研究は「セキュリティ警告の設計者が意図したとおりに警告を活用できない人に対する改善方法を探る」という目的で研究を行っている．それゆえ，設計者の意図したとおりに活用できない人を対象に，意図したとおりに活用しないような状況のもとで評価を行うことが理想である．そういった状況の作り込みについて実験設計としては一定の限界があることについては 5.2 節で述べた．しかし，実験参加者に依頼するタスクには工夫の余地が残されていると考えている．以下の要件を満たしうる仮想タスクを検討することは今後の課題である．

(a) 実験参加者がオンラインで実施可能な実験であること．
(b) 大学研究室が実施する実験タスクとして不自然でないこと．
(c) 実験タスクの実施途中でセキュリティ警告が出現する必然性があること．
(d) セキュリティ警告への対応行動として参加者が判断に悩む，または非推奨行動をとる傾向にある状況であること．
(e) 警告の出現を実験の一部だと確定的に判断できないこと．

また同様の理由から，実験参加者をその属性に応じて事前に振り分けて実験を割り当てることが望ましいと考える．今回の事後アンケートの項目 (6)，(7) から，IT 専門知識の有無は警告への対応行動に対する影響が少ないが，衝動性については対応行動の有無に差が生じることが明らかになった．また非推奨行動をとった 11 名の参加者についてインターネット利用歴と利用頻度をそれぞれ 4 段階で調査をしたところ，利用歴が長く，また利用頻度が高いほど非推奨行動を選択する傾向が見られることも分かった．今回の実験では，実験参加者の属性や知識・経験を実験実施前に調査せず，システムで無作為に各実験条件に割り当てたが，今後は警告設計者の意図したとおりに警告を活用できない人・しない人の特性を明らかにしていき，その特性ごとに警告の改善方法を模索することも必要だと考えている．またこの点から，セキュリティ警告も利用者の特性にあわせ Adaptive かつ Personalize されていく必要性もあるかと考えている．これらの点についても検討を進め，今後の評価実験に反映していきたいと考えている．

(2) かわいい効果に関する情報提示方法：今回の実験では，既存の警告における注意喚起マーク部分をかわいい画像に変更しただけであり，かわいい効果の情報提示法としては抑制的であったと考えている．このような設計にした理由は，他の警告条件との比較検証のためである．しかし，かわいい効果を発生させる提示法としてこれが最善であるかは不明なままである．既存の警告インタフェースの設計でも様々な工夫がなされていることは 2 章で述べたが，かわいい効果の提示法についても工夫の余地があると考えている．また，かわいい効果を発生させるための素材についても何が適切か，また利用者ごとに Personalize 可能にすることが望ましいか，などについても検討を進めていく．

(3) 他の行動誘引効果：ある刺激とそれによって引き起こされる人間の行動に関する知見について，特に「何かの刺激・きっかけを与えると，人は思わずある種の行動をとってしまう」といった誘引効果について継続して調査を行い，それをセキュリティ警告の問題改善に応用できないかについて検討を進める．

(4) 他の言語・文化圏に属する実験参加者による評価：今回の実験における実験参加者は，未確認であるものの，ほぼすべての参加者が日本人であると推測する．そう推測する理由は，依頼タスクやアンケート内容ならびに実験システムにおける表記がすべて日本語であったためである．このことから，今回の実験結果は特定の言語を理解する母集団によるものといわざるを得ない．また文献 [21] の調査によるとアニメーション風の図による警告を「子供っぽい（*childish*）」とネガティブにとらえる人もいた，という結果が得られている．これと同様の印象を “かわいい” イラストに対して持つ人もいるであろう．他の言語圏または文化圏に属する実験参加者による “かわいい効果” の評価を行い，今回の評価結果と比較考察を行うことも今後の課題として検討する．

6. おわりに

ユーザが ICT 技術を使う中で情報セキュリティの脅威に直面する可能性がある．その際，脅威への遭遇を回避させる機能の 1 つとして「セキュリティ警告」がある．しかし，脅威の誤検知や利用状況，知識不足などの理由からセキュリティ警告が有効に活用されていない現状がある．この問題に対し，本研究ではかわいい効果による行動誘引効果をセキュリティ警告の問題改善に応用することを提案した．

クラウドソーシングを通じて 84 名の実験参加者を募り，

© 2020 Information Processing Society of Japan

オンラインアンケートのタスクを依頼する形で評価実験を実施した．評価方法は，警告内に提示される画像が異なる3種類の警告を用意し，各実験参加者はそのうちの1種の警告に1回だけ遭遇する仕組みとして行った．その警告遭遇の際に警告に対する参加者の反応を実験システムを通じて測定するとともに，タスク終了後には，被験者の属性と実験に関する追加調査を事後アンケートとして行った．

実験の結果，警告に対する反応時間ならびに対応行動については，かわいい画像を提示したものも含む3種の警告間で有意な差は見られなかった．しかし，警告内の警告文理解については，かわいい効果のない画像による警告よりもかわいい効果が期待される画像を提示した警告の方が理解率が有意に高いという結果が得られた．

今後は，評価実験において望ましくない影響を与えうる要因を少なくする実験設計・依頼タスクについてさらなる検討を進めるとともに，かわいい効果の提示方法の工夫による効果の差に関する調査も行う．また警告を活用できない利用者の属性調査を行い，その結果に基づくセキュリティ警告の改善手法，さらには利用者に応じて Adaptive または Personalize されるセキュリティ警告についても検討を進めていく．

参考文献

[1] 皆川 諒，高田哲司：馴化を抑制しうる新たなセキュリティ警告の探求：かわいいとその付加刺激の効果に関する評価，コンピュータセキュリティシンポジウム 2017（CSS2017）(2017).

[2] Minakawa, R. and Takada, T.: Exploring alternative security warning dialog for attracting user attention: Evaluation of “Kawaii” effect and its additional stimulus combination, *Proc. iiWAS'17*, pp.582–586 (2017).

[3] Yahoo Japan：Yahoo クラウドソーシング，入手先 ⟨https://crowdsourcing.yahoo.co.jp/⟩（参照 2018-08-08）.

[4] 小倉加奈代：ユーザのフィッシングサイト回避能力と心理特性との関係性の検討，情報処理学会研究発表会 2017-SPT-23 (2017).

[5] Bauer, L., Lillo, C.B., Cranor, L. and Gragkaki, E.: Warning Design Guidelines, CMU-CyLab-13-002 (2018), available from ⟨https://www.cylab.cmu.edu/_files/pdfs/tech_reports/CMUCyLab13002.pdf⟩.

[6] 小橋眞理子，井田政則：改訂日本語版 BIS-11 の作成—信頼性と妥当性の検討，立正大学心理学研究年報，Vol.4, pp.53–61 (2013).

[7] Cranor, L.F.: A framework for reasoning about the human in the loop, *1st Conf. Usability, Psychology, and Security* (*UPSEC'08*) (2008).

[8] Sunshine, J., Egelman, S., Almuhimedi, H., Atri, N. and Cranor, L.F.: Crying Wolf: An Empirical Study for SSL Warning Effectiveness, *USENIX Security Symp.*, pp.399–416 (2009).

[9] Felt, A.P., Ainsile, A., Reeder, R.W., Consolvo, S., Thyagaraja, S., Bettes, A., Harris, H. and Grimes, J.: Improving SSL Warnings: Comprehension and Adherence, *CHI'15*, pp.2893–2902 (2015).

[10] 山崎知嗣：偽警告に騙されないで!～巧妙化する手口とその対策～（情報セキュリティ EXPO 2018），情報処理推進機構 (2018), 入手先 ⟨https://www.ipa.go.jp/files/000066767.pdf⟩（参照 2018-08-08）.

[11] Bravo-Lillo, C., Komanduri, S., Cranor, L.F., Reeder, R.W., Sleeper, M., Downs, J. and Schechter, S.: Your attention please: Designing security-decision UIs to make genuine risks harder to ignore, *Proc. SOUPS'13* (2013).

[12] Bravo-Lillo, C., Cranor, L.F., Komanduri, S., Schechter, S. and Sleeper, M.: Harder to Ignore? Revisiting Pop-Up Fatigue and Approaches to Prevent It, *Proc. SOUPS'14* (2014).

[13] Yang, W., Xiong, A., Chen, J., Proctor, R.W. and Li, N.: Use of Phishing Training to Improve Security Warning Compliance: Evidence from a Field Experiment, *Proc. Hot Topics in Science of Security* (*HoTSoS*), pp.52–61 (2017).

[14] Krol, K., Moroz, M. and Sasse, M.A.: Don't Work. Can't Work? Why It's Time to Rethink Security Warnings, *Proc. Risk and Security of Internet and Systems* (*CRiSIS*) (2012).

[15] Weinberger, J. and Felt, A.P.: A Week to Remember: The Impact of Browser Warning Storage Policies, *Proc. SOUPS'16* (2016).

[16] Jenkins, J.L., Anderson, B.B., Vance, A., Kirwan, C.B. and Eargle, D.: More Harm Than Good? How Messages That Interrupt Can Make Us Vulnerable, *Journal of Information Systems Research*, Vol.4, No.4, pp.880–895 (2016).

[17] Anderson, B.B., Kirwan, C.B., Jenkins, J.L., Eargle, D., Howard, S. and Vance, A.: How Polymorphic Warnings Reduce Habituation in the Brain: Insights from an fmri study, *CHI'15*, pp.2883–2892 (2015).

[18] 藤原康宏，村山優子：コンピュータ利用時の不快感を利用した警告インタフェースの提案，情報処理学会論文誌，Vol.52, No.1, pp.77–89 (2011).

[19] Nittono, H., Fukushima, M., Yano, A. and Moriya, H.: The power of kawaii: Viewing cute images promotes a careful behavior and narrows attentional focus, *PLoS ONE*, Vol.7, No.9, e46362 (2012).

[20] Balebako, R., Jung, J., Lu, W., Cranor, L.F. and Nguyen, C.: “little brothers watching you”: Raising awareness of data leaks on smartphones, *Proc. 9th Symp. Usable Privacy and Security, SOUPS'13* (2013).

[21] Raja, F., Hawkey, K., Hsu, S., Wang, K.C. and Beznosov, K.: A brick wall, a locked door, and a bandit: A physical security metaphor for firewall warnings, *Proc. 7th Symp. Usable Privacy and Security, SOUPS'11* (2011).

推薦文

本論文は，PC やブラウザなどで表示されるセキュリティ警告において，「かわいい」キャラクターを表示することによって，警告の効果を向上させる方式を提案し，安全行動への誘導に与える影響を明らかにしている．パーソナライズの一実装として「かわいい」に着目した発想が際立つ．また実験のデザインの細かさや適切な記載，既存研究との差分が明確に書かれていることに加え，分析と考察についても偏りなく公平に書かれているなど，論文完成度も高い．これらのことから，推薦するに相応しいと判断した．

（コンピュータセキュリティシンポジウム 2018
プログラム委員長 吉岡 克成）

© 2020 Information Processing Society of Japan

皆川 諒

2016 年電気通信大学情報理工学部卒業．2018 年同大学大学院情報理工学研究科情報学専攻修了．在学中は認知心理を応用したセキュリティ警告の改善に関する研究に従事，マルウェア解析にも関心がある．現在，株式会社 NTT Communications 勤務．

高田 哲司 （正会員）

2000 年電気通信大学大学院情報システム学研究科情報システム運用学専攻博士後期課程修了．博士（工学）．2003 年ソニーコンピュータサイエンス研究所研究員．2005 年独立行政法人産業技術総合研究所情報技術研究部門研究員．2010 年電気通信大学大学院情報理工学研究科准教授．現在に至る．ユーザブルセキュリティ，個人認証，情報視覚化に興味を持つ．IEEE-CS 会員．

© 2020 Information Processing Society of Japan