

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

皆川 諒¹ 高田 哲司¹

概要: セキュリティ警告の効果を阻害する原因の1つに馴化がある。この阻害要因を抑制する取り組みが研究されているが、警告への注目を回復させるにとどまり、その後の対応行動までは考慮されていない。そこで本研究では、セキュリティ警告に「かわいさ」に基づく視聴覚効果を付与することで、馴化の抑制と安全行動への誘導を試みた。被験者実験の結果、提案する刺激方法に基づく警告は既存警告と比較して統計的に有意な改善をもたらし、馴化を抑制する効果を発揮しうる可能性が示された。この結果をふまえ、セキュリティ警告における馴化抑制と安全行動への誘導の2点について、今後の展望と残されている課題について議論する。

キーワード: セキュリティ警告, 馴化, 行動モデル, UI 設計, 「かわいい」工学

Exploring New Security Warnings to Reduce Habituation: Evaluation of "Kawaii" and its Additional Stimulus Effects

RYO MINAKAWA¹ TETSUJI TAKADA¹

Abstract: Habituation is one of the causes of the disregard for security warnings. Although many researchers focus the attention for security warnings to solve this problem, they don't consider leading users a better action through security warnings. In this paper, we propose a security warning using "kawaii" and its additional stimulus that could lead reducing habituation and better action. From the subject experiment, we found that these methods significantly influence in reducing habituation than existing warning. Finally, we discuss that the proposed methods have the effectiveness which reduces habituation and suggestions for improvement in future works.

Keywords: Security Warning, Habituation, Human Behavior Model, UI Design, "Kawaii" Engineering

1. はじめに

セキュリティ警告は、悪意のあるコンテンツからユーザーを守るセキュリティ防御策の一つである。これは、フォールスポジティブの可能性も考慮されるような、安全性と利便性のトレードオフを実現すべき状況で用いられる。SSL/TLS 証明書エラー、実行ファイルを開くときなど、ユーザーは様々なタイミングでセキュリティ警告と遭遇するが、現在その効果は疑問視されている。Symantec の報告 [1] によれば、インターネットユーザーの 91% がセキュリティ警告と対面したことがあるのに対し、警告の指示に従

うユーザーは 50% に満たない。警告と対面しているユーザーが中間者攻撃やゼロデイ攻撃に晒されようとしている場合、警告を無視したならば、ユーザーは攻撃に気づかずに被害を受ける可能性が高い。よって、警告が推奨する行動である「安全行動」にユーザーを誘導することは重要な役割である。

ユーザーを安全行動へ誘導しうるセキュリティ警告を開発するうえで、「なぜユーザーがセキュリティ警告を無視するのか」という問題を避けて通ることはできない。Cranor の論文 [3] によれば、セキュリティ警告が表示されたとき、ユーザーは図 1 の行動モデルに従って自身が取るべき行動を決定する。

¹ 電気通信大学
The University of Electro-Communications

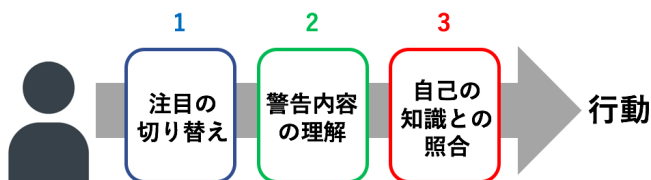


図 1 セキュリティ警告に対するユーザの行動モデル

1 注目の切り替え: ユーザが一時的に作業を中断し、セキュリティ警告に注目を移すまでのステップ
 2 警告内容の理解: 警告文を読み、起こりうるセキュリティ脅威について理解するまでのステップ
 3 自己の知識との照合: ユーザの自身の経験や知識から、有害な影響を与える可能性があるか検討するステップ
 つまり、これらいずれかのステップをユーザが怠った場合、ユーザがセキュリティ警告を無視する可能性が高くなる。先行研究でも、いずれかのステップの改善に着目して警告効果の向上を図っている。Krol ら [2] は『警告内容の理解』に着目し、従来の警告文よりもセキュリティ脅威について具体的に記述した警告文をセキュリティ警告に取り入れ、安全行動への誘導を図った。しかし被験者実験の結果、80%以上のユーザはセキュリティ警告を無視した。さらに、警告を無視したユーザの事後調査から、警告を無視した参加者の 45%は『警告に対する倦怠感があった』と話し、被験者の多くが警告文の内容に目を通していないことを明らかとなった。この結果に対し Krol らは、被験者が過去に遭遇したセキュリティ警告に対する経験の積み重ねが、警告に対する倦怠感として表れているとまとめた。つまり、現状のセキュリティ警告はユーザからの注目を得られておらず、その結果として「警告を無視する」という行動に至っている可能性が高いと考える。なぜなら、「ポップアップ警告を受け付けない」などの理由を合わせると、警告に対する処理の 2 番目である『警告内容の理解』のステップに到達するユーザは、Krol らの実験を基にすると全体の 45%以下しか存在しないためである。よって、ユーザの無視を抑制しうるセキュリティ警告を開発するにあたり、ユーザからの注目を得るための工夫は必要不可欠である。

本論文では、ユーザにとって警告に対する注目が失われている状態を「馴化」と定義する。セキュリティ警告に対して馴化が進行したユーザは警告文に目を通さなくなり、その結果として安全行動をとらなくなる。セキュリティ警告内の警告文を読ませて安全行動へ導くためには、ユーザから失われた警告に対する注目を取り戻す必要があるため、警告への注目を向上させることはセキュリティ警告の効果改善へ貢献できると考える。

上記の理由から、我々はセキュリティ警告の安全行動へ誘導する効果を阻害する要因である馴化に着目した。本稿では、ユーザの注意力を向上させることにより、馴化を抑制しうるセキュリティ警告の開発を試みる。方針として、注目を集めて集中力を上昇させる効果を持つと言われている「かわいい」要素とその付加刺激をセキュリティ警告へ

取り入れ、上記の問題の改善を試みた。実験では、目的を伏せた状態でユーザにセキュリティ警告と遭遇させ、ユーザが意思決定を行うまでの時間と警告の指示に従うかを記録した。最後にこれらの結果をふまえ、提案手法が持つ馴化抑制と安全行動への誘導に関する効果について議論する。

2. 関連研究

2.1 注目の改善に焦点を置いたセキュリティ警告

ユーザからの注目を改善することにより馴化を抑制するセキュリティ警告の先行研究としては、Anderson ら [6] が開発した Polymorphic Warnings がある。Polymorphic Warnings は、警告に揺れや拡大といったアニメーションを付与し、視覚的な刺激を増強することによって注目を集め、その効果を持続させることを試みた。Anderson らは fMRI による脳の活性化領域によって馴化耐性を評価し、結果として既存の警告よりも有意に馴化を抑制する効果があると結論付けた。しかし、ユーザは警告の画像を fMRI の装置内から見たのみで、警告に対して意思決定までは行わなかった。そのため、この手法が安全行動につながるかについては不明瞭のままである。また、Bravo-Lillo ら [4] はユーザに操作を求める警告を設計した。ここでの操作は、警告文の一部をタイピングする操作やマウスでドラックする操作といった、10 秒ほどで完了する簡単な操作を指す。これは、警告文の一領域に対して強制的に注意を向けさせることによって馴化の抑制を図っている。Bravo-Lillo らは 150 秒間に 20 回以上警告と遭遇させることによって馴化させた後、警告文が異なる警告を提示させてユーザがその内容に従うかどうかで評価を行った。結果として、提案手法は警告文の内容に対してユーザを有意に従わせることができ、既存の警告よりも馴化に耐性があるとまとめた。しかし、実験全体で警告に従ったユーザは 14%であり、ユーザを安全行動に誘導できた結論付けるには不十分であると考える。また、Lee ら [7] はログインシールに対してユーザに操作を求める要素を導入した。ここで、ログインシールとは Web サイト上で個人情報を入力する際に、ユーザがあらかじめ登録した画像が出現することを確認することで、正規の Web サイトであるかを判別可能になるフィッシング対策手法の 1 つである。Lee ら [7] は現実的なフィッシングのシナリオに近い形態で実験を行ったがユーザの安全行動率に有意差はなく、むしろ事後アンケートから『ユーザをイライラさせる要素となる』とまとめた。一時的な注目を改善させてもユーザの対応行動に影響を与えることができれば、「馴化による警告の無視を改善した」とは言い難い。

また、ユーザに不快感を与える警告インターフェースを使用させることによって注意を向けようとする研究も報告されている。村山ら [8] は、文字の反転や待ち時間の挿入といった、計算機の利用時にユーザを不快にさせる因子を導入することで注意力を向上させるセキュリティ警告を開発した。実験結果より、被験者の全員が警告の出現を認知で

きたことから注意を引きつける効果があることは確認できたが、警告の指示に従ったユーザの割合(以下、この値を安全行動率と記述する)に有意差はなかった。また、ユーザが不快さをどの程度感じたかについての評価はなく、ユーザにとって不快感が働いているかの指標が曖昧であった。

以上のことから、セキュリティ警告が効果を発揮するために警告への注目を改善することは有益だといえる。しかし、既存手法はユーザの対応行動まで改善しておらず、警告無視の要因である「馴化」の問題を解決できていない。また、ユーザに操作を要求する要素や不快感を与える要素の導入は注意を向けさせることができても、ユーザの安全行動へ結びつける要素としては不適切である可能性がある。

2.2 安全行動への誘導に焦点を置いたセキュリティ警告

安全行動への誘導を目的として、Krolら[2]は警告文の改善を試みた。Krolらは、発生しうるセキュリティ脅威などについて、一般的な警告文よりもより明確に記述したセキュリティ警告を開発し、被験者実験を行った。しかし、安全行動率は実験全体で19%と低く、ユーザが警告を注視した秒数にも有意差は見られなかった。この結果について、ユーザはフォールスポジティブ警告などの存在によって警告疲れを起こし、警告へ注意を払わなくなっていることを事後アンケートから明らかにした。これは、図1の行動モデルにおける「理解」の前段階である「注目」部分を改善する必要があることを示唆している。

また、Feltら[5]は、SSL/TLS警告に関する研究として、一度ユーザが信頼した証明書を保存する期間を調整することによってフォールスポジティブ率を減少させ、安全行動率を向上させることを目指した。ユーザが過去に訪れたWebサイトに再度警告するまでの期間を、セッション単位から一週間に変更することにより、安全行動への誘導率と警告が表示されてからユーザが意思決定を行うまでの時間が改善されることを明らかにした。安全行動率については約14%向上し、この結果からFeltらはフォールスポジティブ警告を取り除くことでユーザの警告疲れを軽減できたと述べた。しかし、この結果はユーザがWebサイトに訪れる度に表示されていたフォールスポジティブ警告の数が減っただけの可能性が高い。セキュリティ警告の役割は真に危険なコンテンツからユーザを安全行動へ導くことだと考えると、フォールスポジティブ率の減少のみで警告疲れを取り除いたと結論付けるのは不十分だといえる。これらの先行研究から、ユーザの警告疲れから表れるセキュリティ警告への注目の低下は、未だに改善すべき課題であると考えられる。これらの点をふまえて、本手法では注意力の向上および安全行動への誘導率を改善するセキュリティ警告を提案する。

3. 提案手法

2章の内容をふまえて、本研究では馴化を抑制するセキュリティ警告に求められる要件を、『利便性を損なわずに、馴

化を抑制する刺激をユーザへ与えること』と定義する。ここで述べる『利便性を損なわない』とは、特別な操作の要求や不快感といった負担をユーザに与えないことを指す。Andersonら[6]の「アニメーション」やBravo-Lilloら[4]らの「ドラック操作による手の動き」が、ここで述べる刺激に該当する。このような視覚的な刺激の増加や視覚以外の刺激を加えることによって、脳が活性化し静的な警告よりも馴化が抑制される[9]。しかし、これらのアプローチは短期間での馴化抑制効果が示されたのみで、長期的な効果と利便性の評価については不明瞭であった。先行研究[4][6]で付与された馴化を抑制するための刺激は1つのみで、時間の経過とともにこれらが効果を発揮し続けるかについては疑問が残る。なぜなら、安全工学の分野においても馴化の防止は課題となっており、異なる知覚に作用する刺激をもたらすことによってこれを防ぐアプローチを取っているため[9]である。そのため、1つの刺激を付与するのではなく、複数の知覚に作用する刺激をセキュリティ警告に付与することが、馴化を抑制するのに最も効果があると我々は考えた。しかし2章で述べたように、タイピングやドラックといったユーザに操作を必要とするタイプの刺激は、利便性を大きく低下させる恐れがある。よって本研究では、ユーザに余計な動作を要求することなく、馴化を抑制することのできる要素を探索する。

これを実現するための1つ目のアプローチとして、我々はセキュリティ警告に「かわいい」要素を導入する。Nittonoら[10]は、簡単なタスクを行う前に小動物のかわいい画像を被験者に見せることで、タスクへの注目を集め、さらに注意力を向上させて慎重に振る舞う効果が得られることを示した。これは、かわいい画像を見たユーザの脳が活性化する可能性を示すとともに、従来手法より警告文を読み、安全行動へ移行する可能性があると考えられる。なぜなら、かわいい画像がセキュリティ警告を無視しようとするユーザの注目を集めた後に、警告と対面することが期待できるからである。さらに、2つ目のアプローチとして、Andersonら[6]が提案したアニメーションを、我々が用意した「かわいい」画像に付与する。アニメーションは知覚刺激を増加させる効果として有用なことはすでに示されているが、アニメーションを付与することによってかわいい画像との相乗効果が生まれる可能性もある。その場合、ただのアニメーション付与よりも馴化の抑制効果が期待できるだろう。最後に、3つ目のアプローチとして犬の鳴き声による聴覚刺激を付与する。Balebakoら[11]は、特定のアプリが位置情報を送信しようとした際、通知機能を用いてユーザに知らせるスマートフォンアプリを開発した。このアプリには水滴音による聴覚刺激を付与しており、半数以上の参加者が音によって認知がしやすくなったと回答した。よって、聴覚刺激はユーザの見落としを防ぐことに効果が大きく、馴化の抑制につながる可能性がある。我々は、聴覚刺激が馴化の抑制に効果があるという仮説を立て、提案警告に鳴き声として組み込んだ。これらの要素を取り入れ、最終的に



図 2 かわいい要素を取り入れたセキュリティ警告

作成したセキュリティ警告を図 2 に示す．使用した「かわいい」要素は，Nittono らの実験で効果が見られた子犬とした．この子犬の画像には図 2 下部に示されている 6 つのフレームによるアニメーションが付与されており，警告が表示されている間は常に外観が変化する．また聴覚刺激としては，警告がユーザの前に現れたとき，犬の鳴き声が発せられる．

4. 被験者による評価実験

3 章で提案したセキュリティ警告が期待される効果を発揮するか検証するために，被験者実験を行った．

4.1 4 つの実験条件

3 章での方針に基づき，我々は 4 つの条件を設定した．被験者は各条件ごとに 4 名，合計 16 名であり，全員が大学生または大学院生であった．表 1 に 4 条件の詳細を示す．

表 1 実験条件

	鳴き声	アニメーション	かわいさによる要素
条件 1			
条件 2	-		
条件 3	-	-	
条件 4	-	-	-

条件 4 は既存の警告と同等の条件である，ベースライン条件として設定した．もし，条件 4 と条件 3 に有意差が認められれば，かわいい要素が馴化の抑制に影響している可能性がある．また，条件 2 と条件 3 の違いはアニメーションの有無であり，この条件の比較によって Anderson ら [6] の提案がかわいい要素に相乗効果を与えるかを見ることが出来る．条件 1 は 2 種類の刺激とかわいい要素を取り入れたセキュリティ警告であり，最も馴化が抑制されると予想している．

4.2 評価方法

本研究では，以下の 2 指標を用いて馴化の抑制度合いを

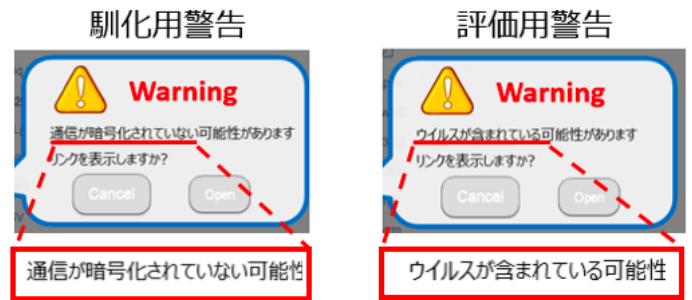


図 3 馴化用の警告 (図左) と評価用の警告 (図右) の警告文

評価する．
意思決定時間

意思決定時間とは，セキュリティ警告がユーザに提示されてから，続行やキャンセルボタンを押すといったユーザが行動するまでにかかる時間を指す．Bravo-Lillo ら [4] の実験結果によれば，ユーザは警告との遭遇を重ねるごとに意思決定時間が短くなり，その時間は 20 回程度の遭遇で 1-2 秒まで低下した．この仮定に基づき，馴化が進行していない状態では警告の意思決定秒数が長くなると予想し，図 3 に示す 2 つの警告を用意して馴化の判定を行った．

図 3 左の警告は，『暗号化通信を行っていない』と告げる警告であり，ログインなどの機密情報を入力しない Web サイトにおいてはセキュリティ脅威となりにくい．よって，機密情報を入力しない Web サイトで繰り返しこの警告を提示するとユーザは次第に馴化し，最終的に意思決定時間は 1-2 秒程度まで低下すると予想する．その一方，図 3 右の警告は，『ウイルスの可能性はある』という内容を告げる警告であり，Web サイトへのアクセスを続行するとセキュリティ脅威となる可能性が高い．しかし，馴化用警告によって馴化が進行した場合，ユーザは警告文の変更に気づかずキャンセルボタンをクリックする可能性がある．このように，ユーザが警告文に目を通さずにキャンセルボタンを押す行動を，本論文では「クリックスルー」と定義する．結果として，馴化が進行したユーザの場合，評価用警告の意思決定時間は馴化用警告のそれとほぼ同じになるだろう．一方，馴化が抑制されて警告文に注目する被験者は，警告文からセキュリティ脅威を認識しようとするため，それに伴い意思決定時間も馴化用警告より長くなるだろう．さらに本実験では，実験条件ごとに被験者の意思決定時間に有意差があるかを見ると同時に，k-means 法による教師なし学習で意思決定時間を 2 クラス分類した際，評価用警告の意思決定時間が「馴化していないクラス」に分類されるかを見ることによって馴化の評価を行う．

安全行動

前述の通り，特定の状況を除いて馴化用警告はクリックスルーをしてもセキュリティ脅威とはならず，その反面，評価用警告はクリックスルーを許すとセキュリティ脅威となる可能性が高い．よって，評価用警告に対して何回安全行動をとることができたかで，馴化の抑制に対する評価を行った．すべての評価用警告に対して安全行動をとるこ

とができれば、馴化による警告効果低減を抑制したといえる。本実験における安全行動は、図 3 右の警告において「Cancel」ボタンをクリックすることである。

4.3 実験方法

評価実験は、被験者の計算機から Web を介して行うオンライン実験形式で行った。この形態をとった理由は、被験者がセキュリティ警告と遭遇したとき、普段とは異なる動作をする可能性を避けるためである。研究室実験で被験者に対して日常的に行わないような操作を要求した場合、被験者は実験であることを意識してセキュリティ警告に注意を向ける可能性がある。

次に、実験手順について示す。実験は、(1) 事前説明、(2) 警告実験、(3) 事後アンケートの 3 段階で実施した。

(1) 事前説明

被験者には実験内容を『一般ユーザがもつセキュリティについての興味・関心についての調査』と説明し、セキュリティ警告の実験であることは伝えなかった。これは、真の実験目的を被験者に話すと、セキュリティ警告に対して普段とは異なる反応を示す可能性があるためである。被験者には、「我々が用意したメールアドレス内にある 20 個のセキュリティ関連の記事を読み、興味を持ったものと持たなかったものに分類するタスクが 1 週間に 1 度、計 3 回分与えられる」と説明した。実験内容について説明を受けた被験者には、本実験で使用する Web サイトに入場するための ID とパスワードが配布され、それが実験で用いる被験者のメールアドレス情報となった。

(2) 警告実験

オンライン実験は、我々が実装したメールクライアントを被験者に操作してもらう形式で行われた。メールクライアントを警告実験に用いた理由は、html メールの利用でリンクを視覚的に簡単に偽装でき、メール送信元のなりすましは現実の攻撃で頻繁に行われているためである。被験者には Web サイトへのリンクを短縮 URL としたものを 20 個メール内に記述し、毎週送信した。記述された 20 個の URL リンクのうち、該当する 8 個の URL リンクをクリックしたときにセキュリティ警告が出現する。被験者は、毎週 8 個の URL のうち 7 個は馴化用の警告と遭遇し、1 個だけ評価用の警告と遭遇した。各週における評価用警告と遭遇するタイミングは、実験 1 週目では 8 回目、2 週目では 7 回目、3 週目は 8 回目である。

また、実験者は URL の詳細については言及せず、『もし見れないような都合があるならば、リンク先を閲覧しなくても良い』とメールの文中に記載した。そのため、被験者は各自の判断で URL のリンク先を閲覧するかどうか決定することができた。

(3) 事後アンケート

3 週間の実験終了後、被験者には本実験に関する本当の目的についての説明を行い、アンケートへの回答を求めた。アンケートでは、被験者に関する属性情報に加えて以下の

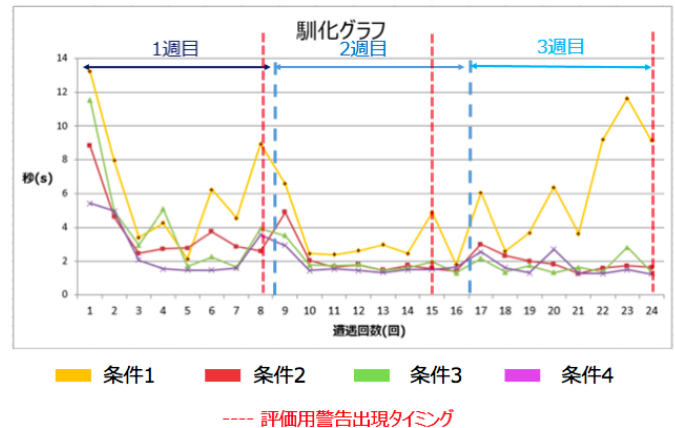


図 4 実験条件ごとに見る意思決定時間の推移

3 項目の質問をした。

実験設計に関する質問:『実験の真の目的に気づいたか』

本手法の効果に関する質問:『警告内容について認知していたか』

利便性に関する質問:『セキュリティ警告に親しみを感じたか』

5. 実験結果

5.1 意思決定時間

実験条件ごとにまとめた意思決定時間の推移を、図 4 に示す。これらの意思決定時間について、条件ごとにマン・ホイットニーの U 検定を用いて有意差検定を行った。1 条件あたりのデータ数は、被験者 4 人 × 警告との遭遇回数 24 回分の 96 である。条件 3 と条件 4 には有意差が認められ ($p < .001$)、「かわいい」要素が意思決定時間に影響を与えていることを示した。一方、条件 2 と条件 3 には有意差が見られず ($p = .41 > .01$)、アニメーションによる視覚刺激は意思決定時間に影響を与えていないという結果となった。最後に、条件 1 と条件 2 には有意差が認められ ($p = .006 < .01$)、鳴き声による聴覚刺激は意思決定時間に影響を与えていることを示した。

また、評価用警告のみの意思決定時間の平均値を図 5 に示す。図 5 中における青の棒グラフが 1 週目の評価用警告における意思決定時間の平均値を示し、同様に赤が 2 週目、緑が 3 週目である。また、図 5 中の赤枠で囲まれている棒グラフが、k-means 法によって『馴化していない意思決定時間』に分類されたグループである。図 5 からわかる通り、条件 1 のグループは時間の経過によって意思決定時間が減少せず、馴化の抑制に効果がある可能性を示している。

図中の赤枠で囲まれている棒グラフが、k-means 法によって『馴化していない意思決定時間』に分類されたグループである。図 5 からわかる通り、条件 1 のグループは時間の経過によって意思決定時間が減少せず、馴化の抑制に効果がある可能性を示している。

5.2 安全行動

評価用警告における条件ごとの安全行動数を、表 2 に示す。条件 1 と 2 では、安全行動をとった被験者は 1 人と

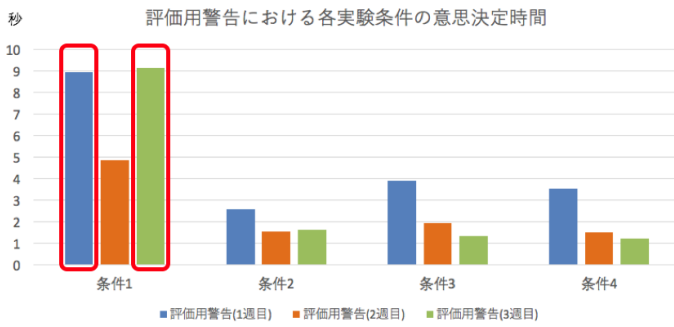


図 5 評価用警告における意思決定時間の平均
赤枠で囲まれたグラフは、k-mean 法で馴化していないと判定された評価用警告

表 2 条件ごとの安全行動数

	安全行動回数内訳	安全行動回数合計	安全行動率
条件 1	3 回:1 人	3 回	25%
条件 2	3 回:1 人	3 回	25%
条件 3	3 回:4 人	12 回	100%
条件 4	3 回:2 人, 1 回:2 人	8 回	66%

どまった。一方、条件 3 の被験者はすべての被験者がすべての評価用警告に安全行動をとり、カイ二乗検定の結果、条件 1 と条件 2 に対して有意差 ($\chi^2(1) = 14.4, p < .001$) が見られた。5.1 節で示した意思決定時間と、安全行動をとった割合について関連は見られず、「かわいさ」と音による刺激が与えた注意力の向上が安全行動へつながっていないという結果となった。「意思決定時間が短い条件の被験者の方が、意思決定時間が長い条件の被験者よりも安全行動をとる」という、意思決定時間と安全行動に不一致が生じた点については、次章で議論する。

6. 考察

6.1 馴化の抑制に有効と思われる刺激に関する考察

まずはじめに、評価実験によって示唆された 3 つの可能性について述べる。既存のセキュリティ警告と遭遇した被験者群 (条件 4) と、かわいい画像を挿入した静的な警告と遭遇した被験者群 (条件 3) の意思決定時間には有意差があった。これは、「かわいい」が馴化を抑制する効果があることを示唆している。しかし、かわいい画像にアニメーションを付与してさらなる視覚刺激の増加を図った被験者群 (条件 2) の意思決定時間には、既存の警告 (条件 4) の意思決定時間との有意差は見られなかった。これは、今回導入した「かわいい」効果は視覚的な刺激であるため、さらに視覚刺激を重ねても大きな効果が得られなかった可能性を示していると同時に、アニメーションが馴化の抑制に大きく働きかけているわけではない可能性を示唆している。その一方、鳴き声による聴覚刺激を警告に付与した被験者群 (条件 1) の意思決定時間は、かわいい画像にアニメーションを付与した被験者群 (条件 2) の意思決定時間

と比較して有意差があった。この結果から、単に刺激の量を増やすよりも刺激の種類を増やす方が馴化の抑制に効果があることを示しており、安全工学分野での知見がセキュリティ警告にも適用できたといえる。

また、既存のセキュリティ警告は他のすべての条件と比較して意思決定時間に有意差が見られた (条件 4 vs 条件 1, 2, 3)。これは、実験 1 週目の 1 回目のセキュリティ警告への意思決定時間を見ても明らかであり、多くのユーザは既存のセキュリティ警告に馴化している可能性を示している。警告効果が低減している現状を考えると、既存のセキュリティ警告は我々が提案した「かわいさ」と音による刺激を取り入れ、失った注目を取り戻す工夫をする必要があるだろう。

6.2 評価指標に関する考察

6.2.1 意思決定時間

本実験では、馴化が抑制されているかを測る指標として「意思決定時間」を用いた。これは、ユーザがセキュリティ警告と対面した際に警告文に注意が向いているかを数値で表すことができる方法である一方、それが真に警告への注目を表す指標で妥当かどうかは疑問であった。そこで、我々は『同じ内容の警告文を含んだセキュリティ警告と遭遇するたびに意思決定時間は短くなるが、馴化が抑制されている状態ならば異なる警告文を提示したときにその変化に気づき、意思決定時間は再び長くなるだろう』という仮説のもと、意思決定時間を馴化の指標として用いた。結果として、意思決定時間の推移は図 4 のようになり、馴化用警告との遭遇を繰り返すことによって意思決定時間が短くなる現象が全ての実験条件を通して見られた。また、馴化を抑制しうる要素を入れた条件では、評価用警告の意思決定時間とその直前にみた馴化用警告の意思決定時間に最大 5 秒の差が確認できた。これらの結果から、我々の意思決定時間における仮説は妥当なものであると考える。

6.2.2 安全行動

提案手法が安全行動へ誘導する効果を持つか調査するために、我々はユーザがセキュリティ警告に対して「Cancel」のボタンをクリックするかどうかをその指標として用いた。しかし、5 章で述べた通り、安全行動と意思決定時間の結果に関連は見られず、提案手法が安全行動へ誘導する効果を持つことを示すことはできなかった。我々は事後アンケートの実施時に「警告内容の変化に気付いたか」という設問を用意し、安全行動と意思決定時間の間に見られる不一致の識別を試みた。先ほどの設問に対して「変化に気づかなかった」と回答した被験者からさらにその理由を引き出すことにより、本実験における安全行動の指標には以下の 2 つの問題があることを明らかにした。

研究室への信頼

1 つ目の問題は、警告内容の変化に気付いたにもかかわらず、評価用警告に対して続行ボタンをクリックした被験者が存在したことである。これに該当する被験者の 1 人

は、『リンクをクリックするか悩んだが、研究室から依頼された実験だったため続行した。普段ならば間違いなく続行しない。』と話した。このような『研究室から依頼された実験だったため安全な実験だと信じた。』という内容を話した被験者は2人存在した。

実験の真の目的への気付き

もう1つの問題は、警告内容の変化に気付かなかったが、評価用警告に対してはキャンセルボタンをクリックした被験者が存在したことである。これに該当する被験者は一人だけ存在し、『セキュリティ警告に対してキャンセルの行動を取れるかどうかを実験の本質だと思った』と話した。実験期間中、被験者は24回セキュリティ警告と遭遇したことを考慮すると、実験の真の目的が警告に対する反応を見ることであることに気付いた被験者も少なからず存在するだろう。他にも警告の真の目的に勤付いた被験者が存在した場合、このバイアスが影響して警告内容の変化に気付き安全行動をとった可能性も考えられる。

6.3 関連研究との比較

6.2節では、馴化の評価指標として意思決定時間を用いることは妥当である可能性を示した。ここで、馴化の評価期間と評価指標について、関連研究との比較結果を表3に示す。

表3 馴化の検証期間および評価指標の比較

	提案手法	Andersonら [6]	Bravo-Lilloら [4]
馴化の検証期間	3週間	5日	150秒
評価指標	意思決定時間 安全行動	fMRI アイトラッキング	警告文の内容に 従うか

馴化は時間をかけて進行するため、可能な限り長期間ユーザをセキュリティ警告と対面させることで、より現実的な環境を反映した実験データを収集することができる。馴化に関する先行研究は短期間の実験であるため、3週間の期間において馴化を抑制効果を示したことは、他の手法よりも長期的な馴化に耐性があることを示唆している。

また、オンラインによる長期間の警告馴化は既存研究にはない試みであり、ここで得られた被験者のデータは、ユーザの反応と馴化の進行度共に現実環境を反映したデータであると考えられる。なぜなら、従来手法が提供するデータは、ユーザがセキュリティ警告の画像を見たときの反応や、警告とは関係のない文章を読んだときの反応であり、セキュリティ警告に対する「意思決定」の反応を含んでいないからである。

さらに、意思決定時間はオンラインによる長期間の被験者実験での測定手法として適しているだろう。馴化の度合いを客観的に測定することは困難であり、先行研究では血中酸素濃度の測定やアイトラッキング [6] といった生体から得られるデータを馴化の指標としてきた。これらの手法は馴化の度合いを数値で表現できる一方、現実の警告遭遇

環境とは異なる状況で測定する必要があるため、ユーザの行動にバイアスがかかる懸念があった。しかし、意思決定時間はソフトウェアにて計測できるため、現実的な環境において警告に対する反応を収集可能である。これは従来研究よりもセキュリティ警告と対面する状況に近い実験環境を再現できる点で、先行研究よりも好ましい実験方法であると考えている。

6.4 利便性に関する考察

我々は、アンケート中で『過去にセキュリティ警告と遭遇したことがある』と回答したユーザに対し、『アンケート中の画像で示すセキュリティ警告に、親しみやすさや使用しやすさを感じていますか』という内容のアンケートを行った。このアンケート項目は、Lee[7]らが用いた利便性の指標に基づいており、この項目の点数が低い場合、ユーザは操作中に負担を感じている可能性が高いことを示す。セキュリティ警告の画像は実験で使用した既存警告と「かわいさ」の要素を取り入れた提案警告の画像を使用し、既存警告については「はい」か「いいえ」の2段階、提案警告については1から5の5段階によって評価を行った。提案警告の評価は、1, 2, 3と回答したユーザについては「親しみやすさを感じていない」として、4または5と回答したユーザについては「親しみやすさを感じている」としてラベル付けを行った。アンケートの結果を、表4に示す。被験者数16人に対して、過去にセキュリティ警告と遭遇したユーザは13人であった。

表4 セキュリティ警告への親しみやすさ、使用しやすさの比較評価

	感じている	感じていない
既存警告	2	11
提案警告	6	7

既存警告と提案警告について、利便性に違いがあるかカイ二乗検定を用いて有意差検定を行ったが、有意差は見られなかった。 $(\chi^2(1) = 2.88889, p = 0.0891 > .05)$ しかし、約18%のユーザしか従来警告に対して親しみやすさを感じていなかった結果に対して、提案警告は約46%が親しみやすさや使いやすさを感じた、と回答していた。利便性の低下はインターフェースの利用自体をユーザから避けられる可能性があると考えられる。なぜなら、フォールスポジティブ警告の多さから、一部のユーザが使用するWebブラウザを変更したという報告事例 [2] があるためである。そのため、従来手法より利便性を向上させつつ馴化を抑制する効果を持つのは本手法の利点であるといえる。

6.5 今後の課題

今後の課題として、実験設計と警告インターフェースの2点について問題点と改善案を述べる。

実験設計の問題点および改善案

6.2.2 節で述べた通り、本実験で用いた「安全行動」の指標は、我々が想定していた通りに働かなかった。これに対する改善案として、以下の2つが考えられる。1つは、警告が発生する状況をより現実の攻撃形態に近づけることで、ユーザが危険な状況におかれていることを認識できるようにすることである。6.2.2 節のアンケート結果によれば、実験者側が用意したアカウントを被験者が使用したとき、被験者は「安全に配慮されたアカウントである」と認識する傾向にあった。特に、本実験で想定したフィッシングメールなどを警戒させる場合、実験者側が用意したアカウントを被験者に使用させても、被験者はフィッシングメールの可能性をほぼ疑わないだろう。なぜなら、実験用に作成された、個人情報を全く取り扱わないようなアカウントを攻撃されるような可能性は非常に少ないと想像できるからである。改良後の被験者実験では、被験者が実験以前から使用しているメールアカウントを利用することで、安全行動の指標に生じるバイアスを抑えることができるだろう。

もう1つは、セキュリティ警告との遭遇回数も最小限にすることである。本実験では、Bravo-Lilloら [4] の実験設計に基づき、ユーザは実験期間中にセキュリティ警告と24回遭遇し、強制的に馴化させられた。しかし、この遭遇回数は現実の環境で起こり得るとは言い難く、かえってユーザが実験の真の目的に気付く要因となっている可能性が高いだろう。図4の意思決定時間の推移を見ると、ユーザは1日に3回セキュリティ警告と遭遇するだけで意思決定時間は大きく減少する。1日に3回という条件も現実の環境で起こる可能性は低いが、将来研究においてセキュリティ警告に対する馴化実験を行う場合、この回数は実験バイアスを最小限にしつつユーザを馴化させるための指標となるだろう。

警告インターフェースの問題点および改善案

我々は、本実験で使用するセキュリティ警告を図3のように設計したが、警告文の内容をユーザが理解できるかについては考慮していなかった。1章の図1に示した行動モデルにもあるように、『警告内容の理解』はユーザを安全行動へ導くために必要不可欠である。しかし、図3に示す通り、我々の設計した警告にはセキュリティ脅威の詳細を知るためのボタンが存在しないため、一部のユーザはなぜ警告が出現したのか理解できない可能性がある。改良後は、Webブラウザのセキュリティ警告と同じく、警告が出現した理由に関して詳細な内容を表示するためのオプションを取り入れることを検討する。

7. おわりに

警告の効果を阻害する原因の1つに馴化があり、警告へ向けるユーザの注意力は馴化の影響によってほぼ失われている。既存研究では、セキュリティ警告にアニメーションや操作を要求する要素を付加することによってユーザが受ける刺激を増加させ、馴化の抑制を図った。しかしこれら

のアプローチは、馴化の抑制効果を示した一方、警告と対峙した後のユーザの対応行動については考慮されていなかった。

そこで本研究では、「かわいさ」による刺激とアニメーションや音による付加刺激を加えることにより、馴化を抑制しつつ安全行動への誘導を可能にするセキュリティ警告を提案した。被験者実験の結果、「かわいさ」が与える刺激とそれに付随させた音による刺激は、意思決定時間に有意差が見られ、利便性を損なうことなく馴化の抑制に効果がある可能性を示唆した。

本論文では、「かわいさ」による条件を主軸において実験を行ったが、「かわいさ」を用いていない外観が風変わりなセキュリティ警告を条件として用意し、「かわいさ」が真に馴化を抑制する効果を持つか検証する予定である。また、被験者がセキュリティ警告に対して実環境に近い状況で意思決定ができるよう、送信メールアドレスや被験者に与えるタスクなど、理想的な実験環境を検討していく所存である。

参考文献

- [1] Symantec: セキュリティ警告, うっかり自社サイトに出ていませんか? ブラウザで表示されるエラーの原因と対処法. 入手先 < https://www.jp.websecurity.symantec.com/welcome/pdf/wp_security_warning.pdf >, (参照 2015-09-18).
- [2] K. Krol, M. Moroz, and M.A. Sasse: Don 't work. can 't work? why it 's time to rethink security warnings, CRISIS, Oct 2012.
- [3] L. F. Cranor: A framework for reasoning about the human in the loop. UPSEC, 8(2008).
- [4] Cristian Bravo-Lillo and Lorrie Cranor and Saranga Komanduri and Stuart Schechter and Manya Sleeper: Harder to Ignore? Revisiting Pop-Up Fatigue and Approaches to Prevent It, SOUPS ' 14 (2014).
- [5] J. Weinberger and A. P. Felt: A Week to Remember: The Impact of Browser Warning Storage Policies, CHI'16 (2016).
- [6] A. Vance, B. Kirwan, D. Bjornn, J. Jenkins and B. B. Anderson: What Do We Really Know about How Habituation to Warnings Occurs Over Time? A Longitudinal fMRI Study of Habituation and Polymorphic Warnings, CHI'17 (2017).
- [7] J. Lee, L. Bauer and M. L. Mazurek: The Effectiveness of Security Images in Internet Banking, IEEE Internet Computing, vol. 19, no. 1, pp. 54-62 (2015).
- [8] 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ーフェースの提案, 情報処理学会論文誌 Vol. 52 No. 1 77-89 (2011).
- [9] 河野 龍太郎. 医療におけるヒューマンエラーなぜ間違えるどう防ぐ. 医学書院, (2014).
- [10] H. Nittono, M. Fukushima, A. Yano and H. Moriya: The power of kawaii: Viewing cute images promotes a careful behavior and narrows attentional focus. PLoS ONE, 7(9):e46362, (2012).
- [11] R. Balebako, J. Jung, W. Lu, L. F. Cranor and C. Nguyen: "little brothers watching you": Raising awareness of data leaks on smartphones. In Proceedings of the Ninth Symposium on Usable Privacy and Security, SOUPS ' 13 (2013).