「かわいい」はセキュリティ警告の効果を改善しうるか？（第２報）～心理効果による安全行動誘引の試み～

皆川 諒[1,†1]　高田 哲司[1]

概要：セキュリティ警告は，ユーザにセキュリティ脅威への遭遇可能性を利用者に伝え，遭遇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」による心理的誘引をセキュリティ警告に導入することを提案する．このアイデアに基づき，セキュリティ警告に「かわいい」効果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しうる，という結果を得た．

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

1. はじめに

「セキュリティ警告」(Security warning) 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挙げられる．

このセキュリティ警告を不要にすることは難しいと考える．本来であれば，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” を自動で実行してくれることが望ましい．しかし以下に挙げる理由から，それは困難である．

- 誤検知・検知不能の問題: 新たな脅威は次々と発見されること，現在存在する脅威を 100%正確に検知するのが困難なこと，不適切なシステム設定により脅威となる可能性がある，などの問題がある．
- 脅威判定の限界: 利用者の利用用途や利用環境に応じて判断が分かれる場合があり，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，現在の状況に対する情報提供と取りうる選択肢を提示し，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 SSL 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．また同様の報告は Krol らの論文 [14] でも報告されている．このような状況は，セキュリティ警告が意図した通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した [1]．具体的には，セキュリティ警告画面に「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改善し，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施結果に有意な差がみられた．これは「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，画像を見た被験者は，

[1] 電気通信大学
The University of Electro-Communications
[†1] 現在，民間企業勤務
Presently with Private Company

注意深く振る舞い，かつ集中力が高まったからであると結論づけている．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った．

しかし，その先行研究 [1] で提案したセキュリティ警告とそれによる評価実験には 2 章で議論するいくつかの問題点が残されており，それゆえに得られた評価結果が「かわいい効果」に起因するものかが曖昧な部分があり，かつ一般化できる結果とは言えない状況にある．

そこで本研究では，セキュリティ警告における問題を「かわいい」が改善しうるか？を明らかにするため，上記の問題をふまえた上であらためて実験設計を行い，セキュリティ警告とそれに対するユーザの行動について ICT 利用者による評価実験をあらためて行った．以降，本論文では 2 章では著者らの先行研究における評価実験の問題点とその改善方法について述べ，3 章では 2 章での議論を基に実施した評価実験の方法と結果について述べる．4 章では，実験結果をふまえて主題に対する結果と，実験を通じて明らかになった問題点や限界について述べる．

2. 先行研究における問題点とその改善方法

著者らによる先行研究 [1] での実験では「セキュリティ警告への馴化」の問題に対して「かわいい」効果を導入したセキュリティ警告が改善効果を示しうるかについて検証を行った．その結果「改善しうる」という結果を導くにいたった．しかし，評価実験及び実験に用いたセキュリティ警告のプロトタイプには以下にあげる問題点が残されており，その結果は「可能性を示唆する」という確度にとどまっていた．本章では，この残されていた問題点とその改善方法について説明する．

2.1 残されていた問題点

残されていた問題点は，以下の 3 つである．

問題 1 実験参加者:

実験参加者は，著者と同じ大学に所属する学部学生および大学院生 16 名であり，性別比は男性:女性 = 7 : 1 であった．また参加者の約半数 (9 名) は情報セキュリティに関する専門教育を受けた学生でもあった．このことから，以下の点で実験参加者の属性はかたよりがあったと言わざるを得ない．

- 特定年齢層 (10,20 歳代) で，かつ男性多数という属性を持つ参加者層
- セキュリティ警告への反応が一般の ICT 利用者とは異なっていた可能性
- 各条件の実験を実施した人数は 4 名 (between-subject study) であり，結果の一般化は困難

図 1 先行研究 [1] におけるセキュリティ警告

問題 2 実験環境:

評価実験のために著者らが用意した環境は，実験専用の Web メールクライアントを実装し，各被験者にそれを一利用者として利用できるようユーザアカウントを発行した．この環境を通じて，電子メールでタスク内容を被験者に通知し，実験タスクを実施させていた．この環境は，被験者が実際に利用している ICT 利用状況と類似してはいるものの，同一ではない．また実験用にユーザアカウントを提供したため，被験者が実際に利用しているアカウントとは異なることから「万が一，なんらかの被害にあっても困ることにはならない」ことが自明であり，その事実が被験者の警告に対する振る舞いに影響を与えた可能性が残されていた．

問題 3 警告プロトタイプ:

先行研究で効果検証に用いたセキュリティ警告を図 1 に示す．このプロトタイプ警告における主たる問題点は，外観が既存のセキュリティ警告と大きく異なる点である．このため，評価によって得られた結果が「外観が大きく変わった」ことによるものなのか，それとも「かわいい」によるものなのかが不明瞭になっている点である．

これらの問題点をふまえ，次にこれらの問題点の改善方法について述べる．

2.2 改善方法

前節で述べた問題に対し，以下のような改良を行なうこととした．

問題 1: 実験参加者の属性におけるかたよりを減らすため，クラウドソーシングを利用して実験参加者を募集することとした．これにより，先行研究による実験参加者群よりも幅広い属性を持つ ICT 利用者による評価が期待できる．

問題 2: 実験環境は，なるべく現実の利用環境に近い状況で評価を行うことが望ましいと考えた．そこで今回の実験では，仮想の Web メールシステムを通じたタスク通知ではなく，実際に実験参加者が利用しているメールアドレ

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

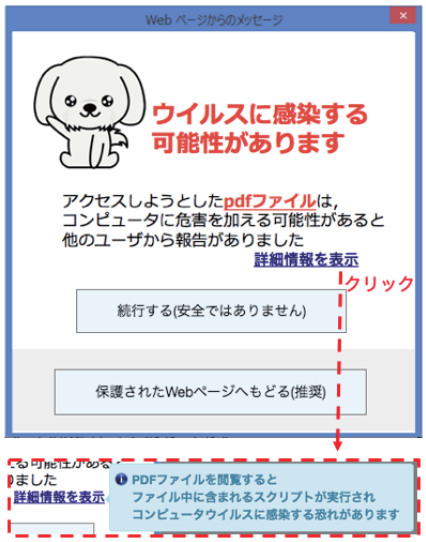

図 3　今回の実験におけるかわいい効果付きセキュリティ警告

スを通じてタスクの依頼をすることとした．ただし，実験自体は Web ブラウザで特定の URL にアクセスさせ，そこから Web アンケートに答えてもらうかたちで実施した．

問題 3: 前述の問題をふまえ，今回の実験で用いたセキュリティ警告は，その外観を既存の警告と同様の「矩形領域」にし，一見した際の外観をほぼ同一化したことである．今回の実験で用いた「かわいい」効果付きセキュリティ警告を図 3 に示す．先行研究で用いたセキュリティ警告図 1 とは外観が大きく異なり，矩形状のダイアログを基本的な外観とし，その中にかわいい画像を表示することで，警告への反応に対する効果を検証することとした．なお本警告のインタフェースに関する設計は，L.Bauer らのガイドライン [5] で推奨されている設計指針に基づいている．推奨された対応を行うボタンを，そうでないボタンよりも大きくする，警告文章は簡潔にしつつ，その詳細を知りたいユーザに対しては詳細情報を提供可能にするなどの設計は，それに従ったものである．

3. 評価実験

前章での議論をもとに，本研究ではクラウドソーシングを利用したオンライン実験を実施した．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．

3.1 実験方法

2.2 節の改善方法をふまえて決定した実験方法について説明する．実験参加者は，Yahoo クラウドソーシング [3] を利用して募集した．参加者に対する参加条件は特に設けず，実験環境として Google Chrome ブラウザが利用できる方とした．募集内容は「実験所要時間は 30 分前後であり，実験完了者には 10 ポイントの報酬を支払う」とした．この募集に対して応募してきた実験参加者に対し，3 種の警告の 1 つにランダムに割り当てた．割り当てはシステムで自動的に行った．この実験は “between-subject design” による実験であり，各被験者は 3 種の警告のうちの 1 つにのみ遭遇した．実験手順の概要は以下の通りである．

(1) 事前説明: 実験内容について，Web ページと書面を閲覧して頂く形で被験者に説明を行った．なお実験への影響を回避するため，本来の実験目的を被験者には明かさず，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ついてアンケート調査する」と説明した．

(2) 実験 (オンラインアンケート): 実験実施側が用意した「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)」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に依頼するタスクとして違和感のないようにするためである．なお，このアンケート回答の途中において，9 問目の回答のあとに図 2 に示す 3 種のセキュリティ警告の 1 つを 1 回出現させた．この警告に対する実験参加者の対応についてシステムを通じて測定した．

(3) 実験に関する事後アンケート: 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アンケート調査を実施した．

(4) 事後説明: 実験の目的が「セキュリティ警告に対するユーザの対応に関する評価」であったこと，また実験結果に影響を与える恐れがあったため，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における内容の詳細については，3.3 節の実験結果とともに説明する．

3.2 評価に用いた警告システム

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．今回の実験では以下の 3 種のセキュリティ警告を実装して検証に用いた (図 2 参照)．

条件 1) 「かわいい画像」によるセキュリティ警告
条件 2) 「かわいさのない画像」によるセキュリティ警告
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ている画像のみとなっている．なお条件 3 は，ベースライン条件として用意した警告であり，条件 1 でかわいい効果による測定を行うことを意図しつつ，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．

3.3 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．

3.3.1 実験参加者について

実験参加者の情報について述べる．本実験では各条件に 50 名の参加者を割り当てる予定で 150 名の募集を行った．しかし，3.1 節の手順通りに実験を完了し，かつ実験データが完全に得られた参加者数は表 **1** の通りで，総数 84 名となった．またその表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専門知識を持つ参加者の人数も示している．なお本論文では，専門知識の有無を「プログラミング言語の学習経験」の有無で判定した．また，表 **2** に全実験参加者の年齢分布を示す．

なんらかの形で実験データが完全に揃わない参加者が 64 名になった理由は，想定外の実験環境で実験を行なった実験参加者がいたためである．3.1 節の手順 (2) では，効果測定を実験システム内で行なっていたが，実装の問題で Google Chrome ブラウザ以外では効果測定が適切に実行できないまま実験を進めることができる状況になっていたためである．なお実験環境は事前説明内に明示していたが，それを読まずに普段利用している Web ブラウザで実験をしてしまったものと推測される．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．

3.3.2 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)

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

	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

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

表 **3** 警告に対する反応

	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)

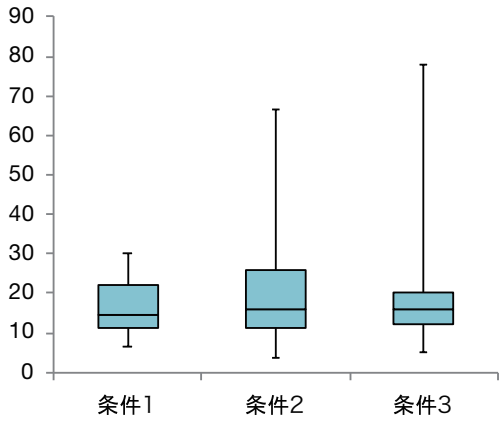

図 **4** 警告への対応時間の分布

表 **4** 警告への反応に要した時間

	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

- 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，推奨行動をとった実験参加者の割合と人数を各条件毎に示している．表 4 は，3 条件の警告に対する実験参加者の対応時間について，中央値，平均値，標準偏差を示している．また図 **4** の箱ひげ図は，各条件における対応時間のデータ分布を示している．

3.3.3 事後アンケートとその結果

実験ならび被験者の属性に関するアンケートを実験終了後に実施した．アンケート内容とその結果を以下に述べる．

(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について，警告から情報を得て，自分で判断し，対応しているかを検証するため，警告に対して行った対応を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，本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，上記のように適切に対応したとは言いがたい，と判定する．回答は，(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 択で行なった．

結果としては，各条件に 1 名づつ計 3 名の実験参加者が自分が行な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，警告をよく見ずに適当に対応をしたものと考えられる．なお 3.1 節の手順 (2) から手順 (3) までにかかる時間は，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想定している．また今回の実験で得られたデータのほとんどは，提示した警告をきちんと見た上で対応された結果であると言える．

(2) 警告内メッセージの理解確認

3.1 節の手順 (2) が完了したあとすぐに，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．回答は正解/不正解として判定し，集計した．

結果は表 3 の最右列に書いてある通りである．警告文をきちんと認識していた参加者の割合は，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なったところ，条件 1-条件 3 のみ “有意差あり”($p = 0.017 < 0.05$) という結果となった．

(3) 警告内の提示画像を “かわいい” と思ったか?

警告内左上に表示した画像について，“かわいい” による効果が意図通りに存在したかについてアンケートを通じて検証した．回答は (かわいい，かわいくない，どちらでもない) の 3 択で実施した．条件 1 の場合は，実験参加者が画像を見て「かわいい」と思ったかを，条件 2 の場合は，「かわいいとは思わなかった」かを検証した．

結果を表 **5** に示す．このことから条件 1 においては，65.5%の参加者が警告内画像を “かわいい” と感じていた．この回答結果を，(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，スコア化したもの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 $p << 0.001$ となった．したがって，条件 1 と条件 2 の警告間では，“かわいい” について異なる効果を発揮していたと言える．

表 **5** かわいい効果の検証結果

	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？

	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 (2)	5 名 (0)
条件 2	22 名	14 名 (2)	8 名 (0)
条件 3	29 名	27 名 (6)	2 名 (0)

(4) 偽警告との疑い

近年，偽の警告画面を表示し，偽対策ソフトウェアのインストールや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，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「偽警告」だとみなし，その判断に基づいて対応を決定した可能性がある．その影響を測るため本調査を行なった．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「偽警告を知っていた」列に書かれた人数が本分析の母集団となる．この結果から偽警告と疑った参加者の割合は，(条件 1, 条件 2, 条件 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になった．

(5) 実験の真の目的に実験途中で気づいたか?

今回の評価実験では，実験への影響を回避するため本来の実験目的を参加者に実験前に説明しなかった．しかし，実験を通じて実験参加者が実験の本来の目的に気づき，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「出現した警告は実験の一部だと思ったか？」を設問の 1 つとし，(1. 一部だと思った，2. 一部だと思わなかった, 3. どちらとも言えない) の 3 択で回答させた．

結果を表 **7** に示す．この表では条件別ならびに推奨行動および非推奨行動をそれぞれ選択した参加者群という観点で集計している．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な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，真の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，いずれの場合

表 7　出現した警告を実験の一部とみなしたか？

	一部だと思った	一部だと思わなかった	どちらとも言えない
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

表 8　対応行動と衝動性

	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

	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

も p 値は $p > 0.05$ となり，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「真の実験目的に気づかれた」ということはなかったと言え，また対応行動が真の目的に気づかれたことにより影響を受けているとも言いがたい，と考える．

(6) 被験者の衝動性

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可能性があると考え，本調査を行なった．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 段階で回答させた．この結果をスコア化して利用している．したがってスコア (S) の値域は，$(8 \leq S \leq 32)$ [*1] となり，スコアが大きいほど「衝動性が高い」ということになる．

結果を表 **8** に示す．この表内の 2 群について t 検定を行なったところ，$p = 0.035 < 0.05$ 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，と言える．

(7) IT 専門知識の有無

本調査では「プログラミング言語の学習経験」を IT 専門知識とみなし，実験参加者にプログラミング言語の学習経験の有無を問うことで回答を得た．各条件において専門知識を有していた実験参加者人数は，表 1 の最右列に記載の通りである．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，各条件毎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%(=(84-11-20)/62) となった．このことから，警告への対応行動が専門知識の有無によって影響を受けているとは言いがたい．

4. 考察

4.1 かわいいによる改善効果について

「“かわいい” はセキュリティ警告の効果を改善しうるか？」という仮説について，今回の実験で得られた 2 種の客観的データから「効果がある」とするデータは得られなかった．

- 対応行動の点で改善効果は見られない:
  フィッシャーの正確確率検定で，3 群 (3 条件) 間に有意差なし
- 対応時間の点でも改善効果は見られない:
  One-way ANOVA による検定で，3 群 (3 条件) 間に有意差なし

ただし，「警告文の理解」に関する調査からは本仮説について示唆的な結果が得られた．それは，かわいい効果付き警告 (条件 1) は警告内の警告文を他の 2 条件よりも有意に認識させていたことが明らかになった．条件 3 をベースラインとして，条件 1-3，条件 2-3 間についてフィッシャーの正確確率検定を行なったところ，2-3 条件間には有意差がないが ($p = 0.0895$)，1-3 条件間には有意差が認められた ($p = 0.0171 < 0.05$)．この結果から，かわいい効果付き警告は警告に注意を向けさせ，警告文を理解させるという点については効果を発揮しうることが明らかになった．

4.2 実験設計とその限界

先行研究 [1] における実験で問題となっていた実験参加者，実験環境，警告プロトタイプのうち，実験環境ならびに警告プロトタイプについては 2.2 節で述べた通りに改善を行なった．実験参加者についても，表 1，表 2 に示す通りクラウドソーシングを通じて実験参加者を募集することで，幅広い属性を持つ母集団により評価が行えたと考えている．また本論文の主題である「かわいい」の効果についても，3.3.3 節のアンケート結果 (3) より，意図通りの効果が一定程度発揮されたものと考えている．

一方，本実験では実験結果に影響を及ぼした可能性のある望ましくない状況が存在していた．それは「理想とする実験環境」と今回の実験環境とを乖離させ，実験結果に望ましくない影響を及ぼしている可能性がある．本実験の理想は「実際の脅威に直面してセキュリティ警告が出現した」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

[*1] 最小値 $1 \times 8 = 8$, 最大値 $4 \times 8 = 32$

警告を「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，各条件で 6 割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いる．したがって，実験結果の解釈も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なっても，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？」と考えることを排除しうる実験設計は困難だ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，評価したい警告画面を「100%偽警告ではない」という確信を持たせた上で評価することは，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，それを行うことは実験として望ましくないことは上記で述べた通りである．

これらのことから言えるのは，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言えるのかを検討していく必要があるという点である．実際の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めるのか？またどの程度まで作り込めれば妥当と言えるのか？は今後の課題の 1 つであると著者らは考えている．

4.3 今後の課題

以下の 2 つを今後の課題とし，本課題に関する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通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なっている．それゆえ，設計者の意図した通りに活用できない人を対象に，意図した通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 4.2 節で述べた．しかし，実験参加者に依頼するタスクにはまだ工夫の余地があると考えている．以下のような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．(1) 実験参加者の計算機で実施可能なオンライン実験であること．(2) 大学研究室が実施する仮の実験タスクとして不自然でないこと．(3) 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 (4) 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．(5) 警告の出現も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通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，その特性毎に警告の改善方法を模索してい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，次の評価実験に反映していきたいと考えている．

(2) かわいい効果の提示方法による効果測定：図 1 と図 3 を見ての通り，先行研究 [1] と今回の実験で提示したかわいい効果付き警告には，その効果の提示方法に大きな差がある．今回の実験では，かわいい効果の提示は抑制的であり，既存の警告の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った．今回は既存の警告との比較検証のためにこのような画面設計としたが，かわいい効果を発揮するためにこの設計が最善であるかは明らかではない．この点について，いくつかの画面設計を検討し，あらためて評価を行う必要があると考えている．また，かわいい画像には何が適切か，また利用者毎に設定可能にする方が望ましいか，などについても検討していく．

(3) 他の誘引効果：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．

5. 関連研究

紙面の都合により詳細は触れないが，セキュリティ警告に関する問題の改善を試みている論文は，以下のカテゴリに分類できる．

- 警告とそれに対応する人間の対応に関するモデル化 [7]
- べからず集 (ガイドライン)[5]
- 警告 User Interface の改良 [11]
- 外部の評価情報の追加 [13]
- 警告内警告文の改善 [8], [9]

- 利用者への提示刺激追加 [17]
- 警告への対応方法の変更 [11], [12], [18]
- 警告の出現タイミングの改良 [15], [16]

本研究で提案している警告は，「利用者への提示刺激追加」に該当する．しかし，既存の研究は新たな刺激を利用者に提示することで警告への対応を利用者に促す (driving) のに対し，本提案はかわいい画像の提示により警告への対応を誘引 (attracting) するという点でアプローチが異なると考えている．

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ことは避けられない．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，知識不足等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画像による誘引効果を用いてセキュリティ警告の問題改善を試みた．

クラウドソーシングを通じて 84 名の実験参加者を募り，オンラインアンケートのタスクを依頼する形で実験を実施した．評価では，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 つに 1 回だけ遭遇する仕組みとし，その際の警告に対する反応について測定を行なった．また，アンケートタスクの後には，被験者の属性と実験に対する意見について事後アンケートを行なった．

実験の結果，警告に対する反応時間ならびに警告によって推奨される行動を選択した推奨行動選択率は 3 種の警告間で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の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，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特性調査を行い，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た Adaptive and/or 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), Oct. 2017.

[2] R.Minakawa, T.Takada, Exploring alternative security warning dialog for attracting user attention: evaluation of “Kawaii” effect and its additional stimulus combination, Proc. of 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] L. Bauer, C.B.Lillo, L. Cranor and E.Gragkaki, 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 第 4 号, pp.53–61, 2013.

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

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

[9] A.P.Felt, A.Ainsile, R.W.Reeder, S.Consolvo, S.Thyagaraja, A.Bettes, H.Harris, and J.Grimes: 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] C.Bravo-Lillo, S.Komanduri, L.F.Cranor, R.W.Reeder, M.Sleeper, J.Downs, S.Schechter: Your attention please: designing security-decision UIs to make genuine risks harder to ignore, Proc. of SOUPS’13, (2013).

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

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

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

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

[16] J.L.Jenkins, B.B.Anderson, A.Vance, C.B.Kirwan, and D.Eargle, More Harm Than Good? How Messages That Interrupt Can Make Us Vulnerable, In Journal of Information Systems Research, Vol.4, Issue.4, pp.880–895, (2016).

[17] B.B.Anderson, C.B.Kirwan, J.L.Jenkins, D.Eargle, S.Howard, and A.Vance, 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 77-89 (2011).

[19] 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).