UNIVERSITÄT BONN

Zur Messbarkeit von IT-Sicherheitsbewusstsein

Ein nutzerverhaltensbasierter Ansatz

Dissertation

zur Erlangung des akademischen Grades
Doctor Rerum Naturalium (Dr. rer. nat.)

vorgelegt von
Arnold Sykosch
geboren in Aachen

vorgelegt an der
Rheinischen Friedrich-Wilhelms-Universität Bonn
Mathematisch-Naturwissenschaftliche Fakultät

im Promotionsfach
Informatik

Bonn, Januar 2022

Angefertigt mit Genehmigung
der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter:	Prof. Dr. Michael Meier Universität Bonn
2. Gutachter:	Prof. Dr. Matthew Smith Universität Bonn
Tag der Promotion:	19. Januar 2022
Erscheinungsjahr:	2022

Danksagung

Ich bin den nachfolgenden Personen zu besonderem Dank verpflichtet. Zuerst möchte ich meinem Doktorvater Michael Meier für seine Ruhe und Zuversicht in den Momenten danken, in denen ich sie am meisten gebraucht habe. Darüber hinaus möchte ich mich bei dem Team des Forschungsprojekts ITS.APT bedanken. Insbesondere Matthias Wübbeling, der mir stets die Gelegenheit für leidenschaftliche und konstruktive Diskussionen bot, Christian Doll für seine Einsatzbereitschaft und Ausdauer bei der Fehlersuche in der Softwareentwicklung und Armin Will, ohne dessen passionierten Einsatz das Projekt nicht zustande gekommen wäre. Ich danke Matthias Wübbeling, Felix Boes und Sophie Jenke für das Lektorat dieser Arbeit. Insbesondere dankbar bin ich meiner Frau Helena und meiner Familie für ihre Geduld mit mir und ihren Glauben an mich.

Kurzfassung

Jedes IT-System ist ein soziotechnisches Konstrukt. Als solches ist es untrennbar mit dem Menschen als Nutzer verknüpft. Auch die Sicherheit dieser Systeme ist mit ihren Nutzern verknüpft. Aus diesem Grund raten einschlägige Handlungsempfehlungen zur Durchführung regelmäßiger Schulungen der Nutzer zur Verbesserung des IT-Sicherheitsbewusstseins. Die Auswirkungen dieser Schulungen werden jedoch nur selten systematisch quantifiziert.

Die vorliegende Arbeit geht der Frage nach, wie bisher zur Quantifizierung von IT-Sicherheitsbewusstsein eingesetzte Verfahren verbessert werden können. Dazu werden zunächst Arbeiten auf diesem Forschungsgebiet analysiert. Es kann herausgearbeitet werden, dass die Arbeiten zum Thema IT-Sicherheitsbewusstsein zumindest motivatorisch von der Handlungsrelevanz dieses Konzepts getrieben werden. Eine Analyse verbreiteter Methoden auf diesem Gebiet zeigt insbesondere Schwächen bei der Erfassung von Nutzerverhalten.

Basierend auf diesem Ergebnis wird das Konzept der *artefaktbasierten IT-Sicherheitsbewusstseinsmessung* entwickelt. Dieses Konzept erweitert den Ansatz der Feldexperimente auf dem Bereich der Forschung zu Phishing-E-Mails. Die relevanten ethischen und juristischen Rahmenbedingungen werden hergeleitet und ausgewertet. Anhand dieser Auswertung wird ein Werkzeug entworfen und implementiert, das die ethische und juristisch unbedenkliche Umsetzung der *artefaktbasierten IT-Sicherheitsbewusstseinsmessung* erlaubt. Dieses wird anschließend in einem Feldexperiment erprobt.

Mit den so gewonnenen Daten lässt sich demonstrieren, dass die artefaktbasierte IT-Sicherheitsbewusstseinsmessung das Verhalten der Nutzer valide erfassen kann. Darüber hinaus erlauben die im Rahmen dieser Arbeit erhobenen Daten eine Reliabilitätsbewertung der artefaktbasierten IT-Sicherheitsbewusstseins-

messung. Diese ist insbesondere auf Phishing-Experimente übertragbar und damit auch in diesem Forschungsbereich relevant. Weiter lässt sich zeigen, dass die Motivation der Probanden allein keinen nachweisbar positiven Effekt auf IT-sicherheitsbewusstes Verhalten hat. Wird ein Maßtyp gewählt, der die Abhängigkeit der Sicherheit der IT-Systeme vom Nutzerverhalten realitätsnaher als bisherige Ansätze modelliert, so lässt sich ein Effekt aufzeigen, der darauf hindeutet, dass eine Schulung der Mitarbeiter der Sicherheit der gesamten IT-Infrastruktur abträglich sein kann.

Inhaltsverzeichnis

1 Einleitung

IT-Infrastrukturen sind soziotechnische Konstrukte und als solche in gesellschaftliche, unternehmerische und politische Strukturen eingebettet [Eck13, 1.1 Grundlegende Begriffe]. Viel stärker noch sind IT-Systeme mit dem Nutzer als Bediener verknüpft. Sie befinden sich in steter Wechselwirkung miteinander. So steht der Nutzer auch in Beziehung mit der Sicherheit dieser IT-Systeme und damit der IT-Infrastruktur, in der diese IT-Systeme eingebettet sind. Das Verhalten des Nutzers kann die Sicherheit der IT-Systeme maßgeblich beeinflussen.

Zudem versuchen Angreife angesichts stetig steigender technischer Sicherheitsanstrengungen den Nutzer in ihr Vorhaben einzubeziehen. Insbesondere werden Nutzer vom Angreifer häufig direkt als Schnittstelle zur IT-Infrastruktur kontaktiert. So nennt der *Verizon Data Breach Report 2020* „Phishing“ als zweithäufigste Angriffstechnik unter allen an Verizon gemeldeten Sicherheitsvorfällen des Jahres 2019 [VER20, Figure 12]. In den Fällen, in denen ein Vertraulichkeitsverlust [Eck13, 1.2 Schutzziele] verifiziert werden konnte, ist das sogar die am häufigsten eingesetzte Angriffstechnik [VER20, Figure 13]. Das Bundesamt für Sicherheit in der Informationstechnik gibt den jährlichen volkswirtschaftlichen Schaden, der in Deutschland durch Phishing verursacht wird, mit einem „zweistelligen Millionenbetrag“ an [Bun20d].

Diese Situation führt zu der einstimmigen Empfehlung der Schulung des IT-Sicherheitsbewusstseins von Nutzern. Das IT-Sicherheitsbewusstsein gilt als maßgeblicher Einfluss auf das sicherheitsrelevante Verhalten der Nutzer. Die Erfolgskontrolle dieser Maßnahmen wird empfohlen, jedoch zeigen bisher eingesetzte Maßnahmen Schwächen. Das Ziel dieser Arbeit ist es, aktuelle Methoden zur Messung von IT-Sicherheitsbewusstsein zu bewerten und, falls möglich, zu verbessern.

Die folgenden Abschnitte stellen die Motivation dieser Arbeit dar, indem das Spannungsfeld beleuchtet wird. Darauf folgend wird das Forschungsprojekt vorgestellt, in dessen Rahmen ein Großteil der hier dargestellten Arbeiten stattfand. In Abschnitt 1.3 werden der wissenschaftliche Beitrag und die zu beantwortenden Forschungsfragen sowie das methodische Vorgehen detailliert beschrieben. Dieses Kapitel schließt mit der Darstellung des Aufbaus dieser Arbeit.

1.1 Bedeutung des „Faktor Mensch“ für die IT-Sicherheit

Aktuelle Standards und Handlungsweisungen zur Erhöhung der Sicherheit von Organisationen empfehlen oder fordern regelmäßige Schulungen der Mitarbeiter im sichereren Umgang mit der IT-Infrastruktur [WH03; Bun20c; Int13; Sta05]. Dies geschieht in der Annahme, dass die Schulung der Mitarbeiter zu einem Verhalten führt, das der Sicherheit der IT-Infrastruktur zuträglich ist. Die Wirksamkeit dieser Schulungen wird jedoch nicht systematisch geprüft. Das im Rahmen dieser Arbeit entwickelte Verfahren lässt diese systematische Prüfung zu. Besonders problematisch und facettenreich zeigt sich die Situation im Bereich der kritischen Infrastrukturen. In diesem Umfeld wird das Verfahren erprobt. Die folgenden Abschnitte stellen die Themenkomplexe detailliert vor.

1.1.1 Zertifizierte Sicherheit

Einschlägige Standards zur Organisation von IT-Sicherheit in Unternehmen bieten die Grundlage zur Zertifizierung der Umsetzung dieser Standards [WH03; Bun20c; Int13; Sta05]. Ein derartiges Zertifikat gilt als Indikator dafür, wie weit IT-Sicherheit in die Unternehmensprozesse eingeflossen ist, und ist so ein Mittel zum Ausweis von Qualität gegenüber Dritten und damit erstrebenswert. Diese Standards verlangen, sofern sie Nutzer von Informationstechnologie betreffen, die regelmäßige Durchführung von IT-sicherheitsbewusstseinsfördernden Maßnahmen innerhalb des Unternehmens.

Dies trifft insbesondere auf den IT-Grundschutz [Bun20c] des Bundesamts für Sicherheit in der Informationstechnik, kurz BSI, und das *Cybersecurity Framework* des *National Institute of Standards and Technology*, kurz NIST, zu [WH03]. Durch

seine Kompatibilität zu ISO 27001/2 und der Auflage in Deutsch und Englisch ist der IT-Grundschutz auch international anwendbar.

Der Teilbaustein „ORP.3 Sensibilisierung und Schulung“ des Bausteins „Organisation und Personal“ (ORP) des IT-Grundschutz-Kompendiums [Bun20c] behandelt den auch in dieser Arbeit fokussierten Themenkomplex. Motivierend wird ausgeführt: „Bei vielen Sicherheitsvorfällen ist die Nichtbeachtung von Regelungen zwar nicht der alleinige Auslöser des Vorfalls, aber mit ein Grund dafür, dass er auftritt“ [Bun20c, ORP.3 2.1]. Weiter wird die Gefährdungslage damit beschrieben, dass „[...] Sicherheitsvorfälle durch das Personal nicht als solche identifiziert werden und auch Cyber-Angriffe bzw. Angriffsversuche unerkannt bleiben“ [Bun20c, ORP.3 2.5].

Darauf aufbauend wird die folgende Zielstellung formuliert: „Mitarbeiter müssen für relevante Gefährdungen sensibilisiert werden und wissen, wie sich diese auf ihre Institution auswirken können. Ihnen muss bekannt sein, was von ihnen im Hinblick auf Informationssicherheit erwartet wird und wie sie in sicherheitskritischen Situationen reagieren sollen“ [Bun20c, ORP.3 1.1]. Dazu soll ein Programm mit dem Ziel ausgestaltet werden „[...] die Wahrnehmung der Mitarbeiter für Sicherheitsrisiken zu schärfen und ihnen die notwendigen Kenntnisse und Kompetenzen für sicherheitsbewusstes Verhalten zu vermitteln“ [Bun20c, ORP 1.2].

Das IT-Grundschutz-Kompendium stellt konkrete Anforderungen an die Organisation, um dieses Ziel zu erreichen: „Um die Mitarbeiter zu sensibilisieren, SOLLTE ein zielgruppenorientiertes Sensibilisierungs- und Schulungsprogramm erstellt werden. Es SOLLTE regelmäßig überprüft und aktualisiert werden“ [Bun20c, ORP.3.A4]. „Die Lernerfolge im Bereich Informationssicherheit SOLLTEN zielgruppenbezogen gemessen und ausgewertet werden, um festzustellen, inwieweit die in den Sensibilisierungs- und Schulungsprogrammen beschriebenen Ziele erreicht sind. Die Messungen SOLLTEN sowohl quantitative als auch qualitative Aspekte der Sensibilisierungs- und Schulungsprogramme berücksichtigen. Die Ergebnisse SOLLTEN bei der Verbesserung des Sensibilisierungs- und Schulungsangebots in geeigneter Weise einfließen“ [Bun20c, ORP.3.A8]. Weiter wird sogar die „spezielle Schulung von exponierten Personen und Institutionen“ gefordert. „Besonders exponierte Personen SOLLTEN vertiefende Schulungen

in Hinblick auf mögliche Gefährdungen sowie geeignete Verhaltensweisen und Vorsichtsmaßnahmen erhalten“ [Bun20c, ORP.3.A9].

Neben anderen wird die „mangelnde Erfolgskontrolle“ als möglicher Grund einer unwirksamen Aktivität in der Aufarbeitung der Gefährdungslage angegeben [Bun20c, ORP.3 2.3]. Die *Umsetzungshinweise zum IT-Grundschutz* des Jahres 2019 verzeichnen elf unterschiedliche Methoden zur Messung des Lernerfolges [Bun19, ORP3.M8]. Dieses Verzeichnis deckt eine ganze Bandbreite von Methoden ab, von einer einfachen Dokumentation der Durchführung einer Schulung bis hin zu Verhaltensexperimenten; es fehlen jedoch konkrete Umsetzungshinweise. Die Umsetzungshinweise zu diesem Baustein wurden bis 2020 nicht aktualisiert [Bun20b].

Der IT-Grundschutz versteht sich selbst als „Interpretation“ der ISO-Normen 27001 und 27002 [Bun18]. Analog fordert ISO/IEC 27001:2013 im Kapitel 7.3 „Awareness“ [Int13], dass Personen, die unter der Kontrolle einer Organisation Arbeit verrichten:

1. die Informationssicherheitsleitlinie kennen,
2. sich ihrer Rolle in der Informationssicherheit innerhalb der Organisation bewusst sind und
3. sich der Implikationen im Klaren sind, wenn sie den Anforderungen des Informationssicherheitsverwaltungssystem nichts Folge leisten.

Die begleitende Sammlung von Leitlinien im ISO / IEC 27002 empfiehlt als Kontrollmechanismus lediglich die regelmäßige Durchführung von *„trainings“* (zu Deutsch oft Schulungen) [Sta05, S. 8.2.2].

Die Sonderausgabe *Building an Information Technology Security Awareness and Training Program* [WH03], Teil des *Cybersecurity Framework* (Baustein *„Awareness and Training (PR.AT)“*), empfiehlt nach der Umsetzung eines *„Awareness Program“* die Feststellung der Effektivität des Programms. Es werden verschiedene Feedback-Mechanismen vorgeschlagen: Fragebögen, Fokusgruppen- und Individualinterviews sowie unabhängige Audits [WH03, S. 6.2]. Auch werden an dieser Stelle konkrete Metriken vorgeschlagen: ein Rückgang von Sicherheitsvorfällen, eine Steigerung der Anzahl sicherheitsrelevanter Meldungen sowie der Anteil der Nutzer, der Schulungsmaterial wahrgenommen

hat. Insbesondere ist hervorzuheben, dass eine Änderung des Nutzerverhaltens erwartet wird.

Zusammenfassend muss festgehalten werden, dass die versuchte Einflussnahme auf Nutzerverhalten bereits in den einschlägigen Standards verankert ist. Weiter wird regelmäßige Intervention sowie die Kontrolle der Effektivität dieser empfohlen. Diese bieten jedoch wenig Anleitung für eine konkrete Implementierung einer derartigen Intervention oder der Feststellung ihrer Effektivität. Die vorliegende Arbeit schließt diese Lücke.

1.1.2 Kritische Infrastrukturen

Alle Bereiche unserer Gesellschaft unterliegen dem Fortschreiten der Digitalisierung. Dies trifft auch auf die sogenannten *Kritischen Infrastrukturen* zu. Das Bundesamt für Bevölkerungsschutz und Katastrophenhilfe gibt die Definition kritischer Infrastrukturen wie folgt an: „Kritische Infrastrukturen (KRITIS) sind Organisationen oder Einrichtungen mit wichtiger Bedeutung für das staatliche Gemeinwesen, bei deren Ausfall oder Beeinträchtigung nachhaltig wirkende Versorgungsengpässe, erhebliche Störungen der öffentlichen Sicherheit oder andere dramatische Folgen eintreten würden“ [Bun20a].

Die Existenz solch kritischer Infrastrukturen wird nicht nur in Deutschland anerkannt. Der Ursprung der staatlichen Aufmerksamkeit entstammt den USA im Jahre 1996 [Sch06] und ist auch fest in der Europäischen Union verankert. Ihr erhöhter Schutzbedarf wird durch verschiedene Sicherheitsstrategien Letzterer gewürdigt [Theo8].

Kritische Infrastrukturen sind in neun Sektoren verteilt: Energie, Informationstechnik und Telekommunikation, Transport und Verkehr, Gesundheit, Wasser, Ernährung, Finanz- und Versicherungswesen, Staat und Verwaltung sowie Medien und Kultur [Bun20a]. Es ergeben sich Interdependenzen bezüglich der Sicherheit der Unternehmen und Institutionen der verschiedenen Sektoren.

Mit der zunehmenden Digitalisierung sind Kritische Infrastrukturen und damit die gesamte Gesellschaft immer weiter von der Sicherheit der eigenen IT-Infrastruktur abhängig. Es entsteht ein besonderer und wachsender Bedarf an IT-Sicherheit in kritischen Infrastrukturen.

1.2 ITS.APT: IT-Security Awareness Penetration Testing

Die meisten der in dieser Arbeit beschriebenen Tätigkeiten wurden innerhalb des interdisziplinären Forschungsprojekts „IT-Security Awareness Penetration Testing“ (kurz: ITS.APT) durchgeführt.

Das Vorhaben startete am 1. Januar 2015 und endete am 30. Juni 2018. ITS.APT wurde durch das deutsche Bundesministerium für Bildung und Forschung als Forschungsinitiative auf dem Gebiet der „IT-Sicherheit für Kritische Infrastrukturen“ im Rahmen des Förderprogramms „IKT 2020 – Forschung für Innovationen“ [Hel13] gefördert. Die Ziele dieses Projekts waren:

- Die Entwicklung eines Maßes zur Bewertung des IT-Sicherheitsbewusstseins.
- Die Erstellung eines praktikablen Werkzeugs zur effizienten Messung des IT-Sicherheitsbewusstseins.
- Die Erstellung eins rechtlichen Rahmenwerks, das praktikable Anleitungen zum Einsatz der ITS.APT-Lösung aufzeigt.

Das Projektkonsortium bestand aus sechs Partnern, ausgewählt nach ihren Kernkompetenzen, um eine spezifische Rolle innerhalb des Konsortiums wahrzunehmen. Das Universitätsklinikum Schleswig-Holstein (kurz: UKSH) übernahm die Rolle des Anwendungspartners. Es stellt das Testfeld zur Erprobung des Werkzeugs und die Anforderungen eines KRITIS-Betreibers an ein derartiges Werkzeug. Die Arbeitsgruppe Allgemeine Psychologie: Kognition an der Universität Duisburg-Essen (kurz: UDE) war für die Laborstudien verantwortlich. Das in diesem Projekt erforderliche Training wurde durch die ERNW Enno Rey Netzwerke GmbH, in Abstimmung mit UDE, konzipiert und durchgeführt. Das Institut für Informations-, Telekommunikations- und Medienrecht der Universität Münster (kurz: ITM) übernahm die arbeitsrechtliche und haftungsrechtliche Einordnung des Projekts. Das Unabhängige Landeszentrum für den Datenschutz Schleswig-Holstein (kurz: ULD) begleitete das Projekt datenschutzrechtlich.

Die Arbeitsgruppe IT-Sicherheit der Abteilung für Sicherheit und verteilte Systeme des Instituts für Informatik der Rheinischen Friedrich-Wilhelms-Universität Bonn verantwortete die Konzeption, die technische Umsetzung und

Durchführung der Testung im UKSH sowie die abschließende Auswertung. Die vorliegende Arbeit fokussiert eben diesen Beitrag. Nach Abschluss des Forschungsprojekts ITS.APT wurde die Veröffentlichung des entwickelten Werkzeugs durch die Europäische Union im Rahmen des Projekts SPARTA des Programms Horizon 2020 gefördert.

1.3 Wissenschaftlicher Beitrag

Die im Rahmen dieser Arbeit verfolgten Forschungsziele lassen sich unter einer Leitfrage subsumieren: *Wie lassen sich IT-Sicherheitsbewusstseinsmessungen verbessern?* Um diese Leitfrage zu beantworten, wird sie in vier Forschungsfragen zerlegt, die im Folgenden kurz vorgestellt werden. Es sind die Teilfragen, die durch die Darstellungen dieser Arbeit leiten. In den Kapiteln dieser Arbeit werden sie entsprechend aufgegriffen.

Forschungsfrage 1: *Was sind Stärken und Schwächen bestehender Ansätze zur IT-Sicherheitsbewusstseinsmessung?*

Die erste Forschungsfrage dient der Aufdeckung von Verbesserungspotenzial bisher eingesetzter Verfahren zur Messung von IT-Sicherheitsbewusstsein. Dazu wird zunächst definiert, was IT-Sicherheitsbewusstsein als Gegenstand der Messung ist. Zu diesem Zweck werden Analysen möglicher Definitionen von IT-Sicherheitsbewusstsein auf ihre Gemeinsamkeiten hin geprüft. Der Konsens dieser Analysen liegt in einem IT-Sicherheitsbewusstseins-Begriff mit Fokus auf der Relevanz im Entscheidungsfindungsprozess in IT-sicherheitsrelevanten Situationen.

Bestehende Ansätze zur IT-Sicherheitsbewusstseinsmessung werden anschließend nach ihrem Typ kategorisiert und bezüglich etablierter Gütekriterien für Messinstrumente bewertet. Dabei zeigen sich verhaltenserfassende Feldexperimente als besonders valide. Jedoch ist ihr Einsatz bisher lediglich auf die Erfassung von Nutzerverhalten in Bezug auf Phishing-E-Mails beschränkt. Damit wird, verglichen mit den anderen Methoden, nicht die gesamte relevante Bandbreite IT-sicherheitsrelevanten Verhaltens erfasst. Auch ist die direkte Erfassung von Verhalten, entgegen fragebogenbasierten Studien, oft mit hohem Ressourceneinsatz verbunden.

Forschungsfrage 2: *Wie lassen sich die Stärken existierender Ansätze in einem neuen Ansatz kombinieren?*

Auf Basis der Arbeitsergebnisse zur Beantwortung der ersten Forschungsfrage wird eine generische Testmethode hergeleitet, welche die Erfassung einer größeren Bandbreite von IT-sicherheitsrelevantem Verhalten ermöglicht und dabei eine vollständige Automatisierung des Ablaufes zulässt. Dabei wird das Konzept der Artefakte als Reizgeber in einer Testsituation entwickelt.

Forschungsfrage 3: *Lässt sich ein konsolidiertes Maß für IT-Sicherheitsbewusstsein konstruieren?*

Auf Grundlage der durch die Testmethode erfassten Daten wird ein Maß für IT-Sicherheitsbewusstsein entwickelt. Das entwickelte Maß konsolidiert dabei bisher eingesetzte Maße und Ansätze, um das erfasste Verhalten zu interpretieren und dieser Interpretation einen Zahlenwert zuzuweisen.

Forschungsfrage 4: *Wie lassen sich artefaktbasierte IT-Sicherheitsbewusstseinsmessungen ethisch und rechtskonform umsetzen?*

Die Antwort auf die vierte Forschungsfrage setzt sich aus zwei Teilen zusammen:

1. Der Erfassung und Auswertung ethischer und juristischer Rahmenbedingungen.
2. Der Konzeption eines Werkzeuges, das eine Messung des IT-Sicherheitsbewusstseins unter diesen Rahmenbedingungen ermöglicht.

Bei der Quantifizierung von IT-Sicherheitsbewusstsein handelt es sich um ein Humanexperiment. Entsprechend stellt sich die Frage nach der Anwendbarkeit des zuvor hergeleiteten Ansatzes. Diesem können insbesondere ethische und juristische Rahmenbedingen entgegenstehen. Um die ethische Anwendbarkeit des Ansatzes zu gewährleisten, wird eine Selbstbewertung durchgeführt. Aus dieser werden *Anforderungen* und *Bedingungen* abgeleitet. Anforderungen sind formulierte funktionale Ansprüche an die konkrete Umsetzung der Messung. Bedingungen wirken sich beschränkend auf die Umsetzung aus. Die so erfassten Bedingungen werden um die Ansprüche angereichert, die sich aus den im Rahmen des Projekts ITS.APT erarbeiteten Handlungsanweisungen ergeben.

Anhand der erfassen Anforderungen und Bedingungen wird ein Werkzeug konzipiert, das in der Lage ist, das IT-Sicherheitsbewusstsein nach dem Prinzip, das als Antwort auf Forschungsfrage 2 gefunden wurde, zu messen. Im Rahmen eines Feldexperiments wird die Anwendbarkeit des Prinzips und des Werkzeugs belegt und damit die Antwort auf die Forschungsfragen 2 bis 4 validiert.

Als Beitrag zum Stand der Wissenschaft ergibt sich so eine Generalisierung des Konzepts von Phishing-Tests zu artefaktbasierter IT-Sicherheitsbewusstseinsmessungen, um bessere IT-Sicherheitsbewusstseinsmessungen zu erlauben. Darüber hinaus ergeben sich der Entwurf und die Implementierung eines rechtskonformen Werkzeugs zur Durchführung von artefaktbasierten IT-Sicherheitsbewusstseinsmessungen sowie einer Studie zur Verifikation des generalisierten Prinzips. Die in der Studie genutzten Artefakte und deren Dokumentation (siehe Anhang A) können als Grundlage für weitere Entwicklung und Durchführung von verhaltensbasierten IT-Sicherheitsbewusstseinstests dienen.

Teile der vorliegenden Arbeit basieren auf einem bereits veröffentlichten Konferenzbeitrag:

> Sykosch, Arnold u. a.: „Generalizing the Phishing Principle: Analyzing User Behavior in Response to Controlled Stimuli for IT Security Awareness Assessment“. In: *Proceedings of the 15th International Conference on Availability, Reliability and Security*. ACM, 2020

1.4 Aufbau der Arbeit

Im folgenden Kapitel 2 werden die Stärken und Schwächen bestehender Ansätze zur Quantifizierung von IT-Sicherheitsbewusstsein herausgearbeitet. Dazu wird zunächst ein Begriff für IT-Sicherheitsbewusstsein definiert, womit mögliche Messpunkte identifiziert werden. Danach werden bisherige Arbeiten vorgestellt, kategorisiert und systematisch bewertet.

In Kapitel 3 wird das Prinzip der verhaltensbasierten Messung von IT-Sicherheitsbewusstsein generalisiert. Anschließend wird ein Interpretationsschema für das erfasste Verhalten sowie ein konsolidiertes Maß für IT-Sicherheitsbewusstsein vorgestellt.

Die juristischen und ethischen Rahmenbedingungen werden in Kapitel 4 strukturiert erfasst. Diese Rahmenbedingungen bilden die Grundlage für die Konzeption und Implementierung eines Werkzeuges zur Messung von IT-Sicherheitsbewusstsein in Kapitel 5. Kapitel 6 dokumentiert die Durchführung der Studie zur Verifikation des entwickelten Verfahrens und Werkzeugs. Die Arbeit schließt mit einer Diskussion der Ergebnisse und einem Ausblick auf weitere Arbeiten.

2 Messung von IT-Sicherheitsbewusstsein

In diesem Kapitel wird die erste Forschungsfrage beantwortet und die Stärken und Schwächen bisher entwickelter Verfahren, die zur Messung von IT-Sicherheitsbewusstsein eingesetzt werden, werden herausgearbeitet. Dazu wird zunächst eine Definition für den Begriff des IT-Sicherheitsbewusstseins, den Messgegenstand, abgeleitet. Darauf folgend wird ein Überblick über die bisher in diesem Bereich eingesetzten Messmethoden gegeben. Anschließend werden diese diskutiert und bewertet, damit Verbesserungspotenzial aufgedeckt werden kann. Letztere führen anschließend zur Herleitung der artefaktbasierten Messung von IT-Sicherheitsbewusstsein, die in Kapitel 3 beschrieben wird.

2.1 IT-Sicherheitsbewusstsein als Messgegenstand

Der Begriff „IT-Sicherheitsbewusstsein“ (englisch „*IT security awareness*“, „*information security awareness*“ oder „*information technology security awareness*“, kurz oft „ISA“) findet in der Mitte der 1990er Jahre Einzug in die IT-Sicherheitsfachliteratur. Scheinbar unabhängig voneinander unternehmen Michel Kabay [Kab94] und Kevin McLean [McL95] erste Versuche, Konzepte aus der Psychologie auf die IT-Sicherheit zu übertragen. Dabei setzen die Autoren einen klaren Fokus auf IT-Sicherheitsbewusstsein als einen bestimmenden Faktor für sicherheitsrelevantes Verhalten. In der Fachliteratur zeigt sich seither jedoch keine klare Definition des Begriffs „IT-Sicherheitsbewusstsein“. Da diese Arbeit zum Ziel hat, IT-Sicherheitsbewusstsein zu messen, ist eine Definition dieses Begriffs unumgänglich. Die akademische Publikationslage präsentiert sich uneinheitlich und motivierte bereits mehrere Literaturübersichten zu diesem Thema [HB14; Häu15; Tso+08; Leb+13]. Diese Arbeiten werden in Bezug auf die

Erkenntnisse bezüglich einer Definition des Begriffs IT-Sicherheitsbewusstsein kurz vorgestellt und darauf folgend ausgewertet.

Tsohou u. a. [Tso+08] gehen bei der Literaturaufbereitung nach der Methode der *Grounded Theory* vor. 48 Publikationen zum Thema IT-Sicherheitsbewusstsein werden untersucht. In dieser Arbeit wird darauf hingewiesen, dass die Begriffe *awareness*, *training* und *education* oft vermischt werden, ein signifikanter Anteil der Publikationen diese Bereiche jedoch voneinander abgrenzt:

> *„Most definitions imply that ISA is the bottom level of a security learning pyramid: ISA aims at attracting the attention of all IS[1] users to the security message, making them to understand the importance of information security and their security obligations, training aims at building knowledge and developing the relevant skills and competencies, and education aims at creating expertise.“* [Tso+08]

Darüber hinaus wird die dokumentierte Motivation der Arbeiten analysiert. Zur Verbesserung der IT-Sicherheit soll zumeist eine Änderung im Verhalten der Individuen, hier *„IS users“*, herbeigeführt werden.

Auch Lebek u. a. besprechen Literatur aus dem Bereich der IT-Sicherheitsbewusstseinsforschung. Ihre Forschungsfrage richtet sich nach den Modellen, die in der Literatur der Verhaltensforschung verwendet werden, um das Verhalten von Individuen zu erklären. Im Großteil der untersuchten Literatur (79 %) finden sich vier verhaltenserklärende Modelle: *Theory of Reasoned Action / Planned Behavior*, *General Deterrence Theory*, *General Deterrence Theory* und das *Technology Acceptance Model*. Aus diesen Modellen werden die verhaltenserklärenden Einflüsse entnommen. [Leb+13; Leb+14]

Hänsch & Benenson teilen gängige Definitionen in drei Dimensionen ein: *Wahrnehmung* (englisch *„Perception“*), *Schutz* (englisch *„Protection“*) und *Verhalten* (englisch *„Behavior“*). In dieser Arbeit bezeichnet die Dimension *Wahrnehmung* die Fähigkeit eine Gefahrensituation zu erkennen, die Dimension *Schutz* als das Wissen um mögliche Handlungsoptionen und die Dimension *Verhalten* als das durch den Nutzer gezeigte Verhalten. Als Motivation für die Umsetzung

[1] Die Abkürzung „IS“ in obigem Zitat ist in der entsprechenden Publikation [Tso+08] nicht spezifiziert. Der Gebrauch lässt jedoch darauf schließen, dass es sich um die Bezeichnung für „IT-Infrastruktur“ handelt.

von IT-sicherheitsbewusstseinsfördernden Maßnahmen wird die Reduktion von Sicherheitsvorfällen genannt. Hänsch & Benenson stellen darüber hinaus die Heterogenität der Evaluationsmethoden heraus, die genutzt werden, um IT-Sicherheitsbewusstsein zu quantifizieren. Sie führen diese Heterogenität auf das uneinheitliche Verständnis von IT-Sicherheitsbewusstsein zurück. [HB14]

Häußinger schlägt eine Systematisierung der Definitionen für IT-Sicherheitsbewusstsein nach drei unterschiedlichen Aspekten vor: *Kognition*, *Verhalten* und *Prozess*. Die Begriffe werden konsekutiv geordnet. Definitionen, die den kognitiven Aspekt von IT-Sicherheitsbewusstsein in den Vordergrund stellen, verstehen IT-Sicherheitsbewusstsein als „Geisteszustand“ (englisch „*state of mind*“) und damit als Voraussetzung IT-sicherheitsbewussten Verhaltens. Der zweite Aspekt ist das IT-sicherheitsrelevante *Verhalten* der Nutzer selbst. Definitionen, die unter dem Aspekt *Prozess* zusammengefasst werden, adressieren die Anstrengungen, die unternommen werden, um das Verhalten der Nutzer zu beeinflussen. [Häu15, Study I: Information Security Awareness]

2.1.1 Diskussion

Um zu einer Definition von IT-Sicherheitsbewusstsein zu kommen, werden die Bestandteile oben aufgeführter Arbeiten neu zusammengesetzt. Diese Rekomposition ist in Abbildung 1 dargestellt.

Vor dem Hintergrund dieser Analysen lässt sich zunächst feststellen, dass die Begriffe *IT-Sicherheitsbewusstsein* und *Verhalten* untrennbar miteinander verknüpft sind. Die Definition von Hänsch & Benenson [HB14] sowie Häußinger [Häu15] inkludieren das Verhalten direkt. Lebek u. a. [Leb+13; Leb+14] verstehen unter IT-Sicherheitsbewusstsein verhaltensbestimmende Faktoren. Tsohou u. a. [Tso+08] stellen explizit dar, dass die in der ausgewerteten Literatur dargestellte Motivation eine Veränderung des Verhaltens von Nutzern zur Steigerung der IT-Sicherheit ist. Damit ergibt sich das sicherheitsrelevante Verhalten von Nutzern als integrale Komponente der Rekomposition.

Die durch Hänsch & Benenson [HB14] geformten Begriffe „Wahrnehmung“ und „Schutz“ sowie der von Häußinger [Häu15] verwendete Begriff der „Kognition“ sind kompetenzbasiert. Ein Nutzer soll in die Lage versetzt werden, den an ihn gestellten Anforderungen zu entsprechen. Insbesondere Tsohou u. a. [Tso+08]

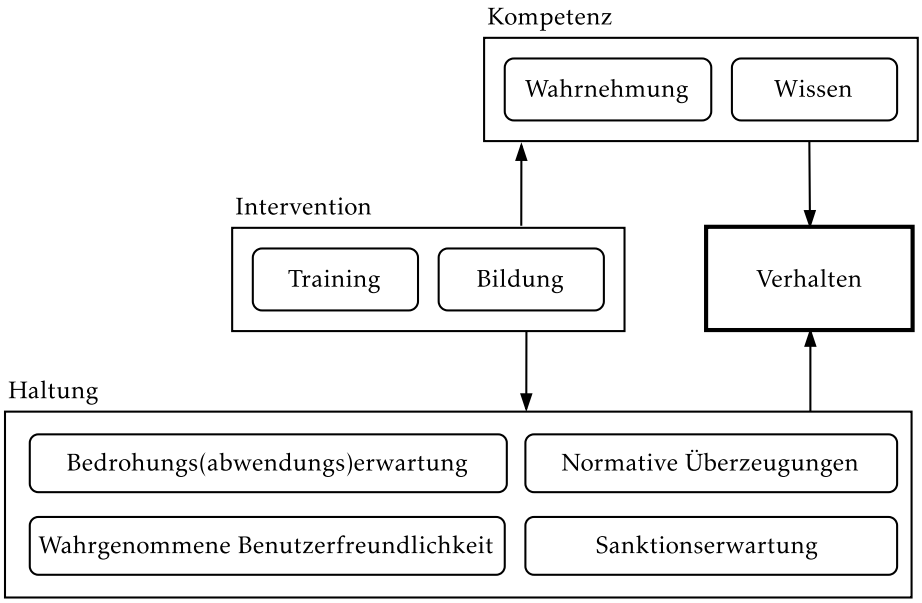


Abbildung 1: *Rekomposition der bestimmenden Begriffe der Definitionen von IT-Sicherheitsbewusstsein. Einflüsse sind als Pfeile dargestellt.*

stellen mit ihrem Verständnis des Begriffs „*awareness*“ heraus, dass der Nutzer diese Anforderungen kennen soll. Dieser Aspekt wird in Abbildung 1 in dem Block „Kompetenz“ entsprechend zusammengefasst.

Häußinger [Häu15] nutzt den Begriff *Prozess*, um die organisatorischen Anstrengungen zusammenzufassen, die das IT-sicherheitsrelevante Verhalten der Nutzer verändern sollen. Tsohou u. a. [Tso+08] differenziert in diesen Anstrengungen durch die Begriffe „*Training*“ und „*Education*“. Diese Anstrengungen stellen eine äußere Einflussnahme dar und werden deshalb in Abbildung 1 unter dem Begriff „Intervention“ subsumiert.

Die durch Lebek u. a. [Leb+13; Leb+14] herausgearbeiteten, modellbasierten Einflussfaktoren adressieren die Haltung eines Individuums. Entgegen dem Begriff der Kompetenz, der Fähigkeiten des Nutzers sich gewünscht zu verhalten, werden hier die individuellen Erwartungshaltungen und Überzeugungen als Teil des IT-Sicherheitsbewusstseins dargestellt. Sie bilden in Abbildung 1 das Konstrukt „Haltung“.

Um festzustellen, welche Konzepte dem IT-Sicherheitsbewusstsein zuzuordnen sind, wird zusätzlich der motivatorische Gedanke herangezogen. Die Forschung in diesem Bereich geht davon aus, dass die IT-Sicherheit vom Verhalten ihrer Nutzer abhängig ist. Das Verhalten der Nutzer kann, im Wunsch einer Erhöhung der IT-Sicherheit, nur bedingt direkt beeinflusst werden. Jedoch kann durch Intervention auf die handlungsbestimmenden Faktoren eingewirkt werden. Die Bestimmungsversuche dieser Faktoren haben die oben dargestellten Konzepte hervorgebracht. Diese werden unter dem Begriff IT-Sicherheitsbewusstsein subsumiert. Abbildung 2 illustriert die resultierende Kette der Abhängigkeiten.

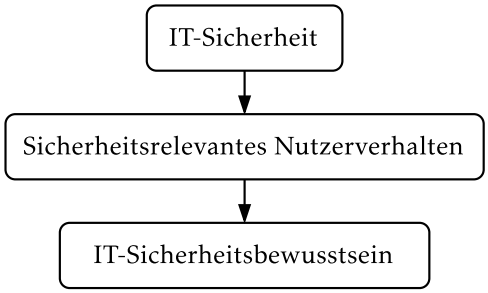


Abbildung 2: *Abhängigkeitskette zwischen IT-Sicherheit und IT-Sicherheitsbewusstsein.*

Es ist festzustellen, dass die Konstrukte *Kompetenz* und *Haltung* Teil des IT-Sicherheitsbewusstseins der Nutzer sind. Das Verhalten selbst oder die Interventionen mit dem Ziel, dieses zu verändern, können durch ihre Abhängigkeiten nicht selbst Teil des IT-Sicherheitsbewusstseins sein. Ob damit alle relevanten Konstrukte des IT-Sicherheitsbewusstseins erfasst sind, kann an dieser Stelle nicht beantwortet werden. Wird das IT-Sicherheitsbewusstsein durch seinen Effekt auf das Verhalten gemessen, so kann nicht davon ausgegangen werden, dass damit das gesamte IT-Sicherheitsbewusstsein erfasst wird. Hier wird lediglich das *effektive* IT-Sicherheitsbewusstsein erfasst. Das effektive IT-Sicherheitsbewusstsein bildet den motivatorischen Gedanken der Forschung in diesem Bereich ab und liegt aus diesem Grund im Fokus dieser Arbeit.

2.2 Bekannte Methoden zur Messung von IT-Sicherheitsbewusstsein

In der Literatur werden vier unterschiedliche Methoden zur Messung von IT-Sicherheitsbewusstsein eingesetzt. Jede Methode nutzt ein *Instrument* zur Vermessung eines *Messgegenstandes* mithilfe eines spezifischen *Maßes* und wird durch einen *Testanwender* durchgeführt. Ist die Testdurchführung personell aufwendig, können mehrere Testanwender zum Einsatz kommen. Diese Methoden werden in diesem Kapitel bewertet und ihre Stärken und Schwächen herausgearbeitet. Dazu werden im nächsten Abschnitt Gütekriterien vorgestellt, anhand derer die Messmethoden bewertet werden können. Diese werden dann in den folgenden Abschnitten dargestellt und, anhand von Beispielen für ihre Anwendung im Kontext der IT-Sicherheitsbewusstseinsmessung, diskutiert und bewertet. Der Abschnitt schließt mit einem Fazit, in dem die Bewertung zusammengefasst wird und Verbesserungspotenzial aufgezeigt wird.

2.2.1 Etablierte Gütekriterien psychologischer Tests

Als Gütekriterien für eine Methode zur Messung von IT-Sicherheitsbewusstsein kommen vornehmlich die gleichen Gütekriterien infrage, die auch für psychologische Tests angesetzt werden [DB16]. Für psychologische Tests existieren drei Hauptkriterien und sieben Nebenkriterien. Die Hauptkriterien reflektieren die Ansprüche an einen Test, um Schlussfolgerungen aus dessen Ergebnissen zu ermöglichen. Die Nebenkriterien referenzieren anerkannte und favorisierte Eigenschaften eines Tests.

Die drei Hauptkriterien für psychologische Tests nach Döring & Bortz sind:

Objektivität: „Unabhängigkeit des Testergebnisses von der Person des Testanwenders/Testleiters“ [DB16, S. 443]

Reliabilität: „Keine oder geringe Verzerrung des Testwertes durch Messfehler, d. h. der Testwert bildet die wahre Merkmalsausprägung der Testperson sehr präzise ab“ [DB16, S. 444]

Validität: „Der Testwert misst tatsächlich das Merkmal, das er laut Testbezeichnung bzw. Testbeschreibung zu messen beansprucht und primär kein anderes“ [DB16, S. 446]

Die Nebenkriterien für psychologische Tests nach Döring & Bortz sind:

Skalierbarkeit: „Ein Test erfüllt das Kriterium der Skalierung, wenn der Testwert durch eine gültige Verrechnungsvorschrift aus den Test-Items[2] gebildet wird [...]“ [DB16, S. 449]

Normierung: „Ein Test ist normiert, wenn aktuelle Testnormen (durchschnittliche Testergebnisse repräsentativer Vergleichsstichproben) vorliegen, die eine Einordnung individueller Testwerte erlauben (normorientiertes Testen).“ [DB16, S. 449]

Testökonomie: „Ein Test ist ökonomisch wenn er in Relation zum Erkenntnisgewinn

- eine kurze Durchführungszeit beansprucht,
- wenig Material verbraucht,
- einfach zu handhaben ist,
- als Gruppentest durchführbar und
- schnell und bequem auszuwerten ist.“ [DB16, S. 449]

Nützlichkeit: „Ein Test ist nützlich, wenn er ein für Praxis und/oder Forschung relevantes Merkmal misst, für das bislang überhaupt kein Test oder nur ein Test mit beschränkter Testgüte vorlag.“ [DB16, S. 449]

Zumutbarkeit: „Ein Test ist zumutbar, wenn er die Testpersonen in Relation zum Erkenntnisgewinn nicht übermäßig in

- zeitlicher,
- körperlicher und/oder
- psychischer Hinsicht belastet.“ [DB16, S. 449]

Nicht-Verfälschbarkeit: „Ein Test ist nicht-verfälschbar, wenn es Testpersonen kaum gelingen kann, absichtlich ein besonders gutes oder besonders schlechtes Testergebnis zu erzeugen, das nicht als unplausibel oder gefälscht auffällt.“ [DB16, S. 449]

Testfairness: „Ein Test ist fair, wenn er allen Personengruppen, für die er anwendbar sein soll, gleiche Chancen bietet bzw. wenn er keine Testpersonen systematisch aufgrund ihrer ethnischen, soziokulturellen oder geschlechtsspezifischen Gruppenzugehörigkeit benachteiligt.“ [DB16, S. 449]

[2] „Ein Test besteht gewöhnlich aus mehreren unterschiedlich schweren Aufgaben oder Fragen (Test-Items), die die Testperson lösen oder beantworten muss.“ [DB16, S. 461]

Die folgende Bewertung der Messmethoden fokussiert die Hauptkriterien. Die Kriterien Zumutbarkeit, Nicht-Verfälschbarkeit und Testfairness sind nicht hauptsächlich von der Messmethode abhängig, sondern der konkreten Implementierung. Sie werden daher nicht weiter betrachtet. Die Skalierbarkeit, Normierung, Testökonomie und Nützlichkeit haben eine enge Verbindung zur Methode. Die folgend vorgestellten Methoden werden auch auf diese Kriterien hin bewertet.

2.2.2 Befragungen

Die Befragung ist die am häufigsten eingesetzte Methode. Lebek u. a. geben an, dass 50 % der durch sie ausgewerteten Arbeiten quantitative Messmethoden einsetzen [Leb+14]. Hier können zwei Messinstrumente zum Einsatz kommen: Ein Fragebogen oder ein strukturiertes Interview durch einen Gesprächspartner.

Ein Fragebogen ist ein etabliertes Messinstrument der Psychometrie (vgl. [DB16, 10.4.3 Psychometrische Tests]). Die einzelnen Fragen werden *Items* genannt. Es werden immer mehrere Items genutzt, um ein *Konstrukt*, eine nicht direkt beobachtbare Variable, zu messen. Das ist der Tatsache geschuldet, dass ein Item allein ein Konstrukt nicht vollständig abbilden kann. Konstrukte entstammen erklärenden Modellen der Psychologie (siehe dazu auch Abschnitt 2.1). Um die Existenz der Wechselwirkung zu zeigen, wird die Korrelation zwischen den Antworten der jeweiligen Konstrukte gemessen. Als Metrik kommt hier oft der Korrelationskoeffizient oder ein vergleichbares Abhängigkeitsmaß zum Einsatz.

Chan, Woon & Kankanhalli zeigen den Einfluss der Wahrnehmung der Wichtigkeit von IT-Sicherheit unter Kollegen und Vorgesetzten sowie der Selbstwirksamskeitserwartung auf das IT-sicherheitsrelevante Verhalten. Dafür kommt ein Fragebogen als Instrument zum Einsatz; die Metrik ist die Korrelation zwischen den Items der Konstrukte. [CWK05]

Aber auch abweichende Ansätze kommen zum Einsatz. Talib, Clarke & Furnell illustrieren mithilfe eines Fragebogens, dass Individuen, die zuvor eine IT-Sicherheitsbewusstseinsmaßnahme durchliefen, sich selbstbewusster in Bezug auf IT-Sicherheit fühlen als solche, die keine Schulung durchlaufen haben. Dazu werden Items genutzt, die das Verhalten und die Kenntnis verschiedener Begriffe aus dem Bereich der IT-Sicherheit adressieren. [TCF10]

Veseli quantifiziert die Effektivität einer IT-Sicherheitsbewusstseinsmaßnahme. Zu diesem Zweck wird eine Eingruppen-Pretest-Posttest-Studie [SCC02, One-Group Pretest-Posttest Design] implementiert. Die Konstrukte Wissen, Haltung und Verhalten (englisch *Knowledge*, *Attitude*, *Behavior*) werden abgefragt. Als Metrik wird ein Index von 1 bis 100 auf Basis der erwünschten Antworten berechnet. Ein positiver Effekt, entsprechend der Metrik, kann bei den Teilnehmern an einer IT-Sicherheitsbewusstseinsmaßnahme demonstriert werden. [Ves11]

Auch Kruger & Kearney setzten die Konstrukte Wissen, Haltung und Verhalten ein, um das IT-Sicherheitsbewusstsein zu vermessen. Als Metrik kommt die relative Anzahl der Probanden, die mindestens 60 % der gewünschten Antworten gegeben haben, zum Einsatz. [KK06]

Alseadoon u. a. korrelieren Persönlichkeitsmerkmale der Probanden mit Items, die das hypothetische Verhalten erfassen. Der Fragebogen erfasst neben den *Big Five* (Offenheit für Erfahrungen, Gewissenhaftigkeit, Extraversion, Verträglichkeit und Neurotizismus) [MJ92] auch Gefügigkeit (englisch *Submissiveness*). Die Probanden werden angewiesen, für fünf in Bildern dargestellte Phishing-E-Mails anzugeben, wie wahrscheinlich sie auf eine der präsentierten E-Mails antworten würden. Es kann eine hohe Korrelation zwischen der Tendenz auf die Phishing-E-Mails zu antworten und den Persönlichkeitsmerkmalen Gefügsamkeit und Offenheit festgestellt werden. [Als+12]

Die *Objektivität* bei fragebogenbasierten Studien ist, verglichen mit anderen Methoden, einfach zu gewährleisten. Dies ist in der stark eingeschränkten Interaktion zwischen dem Testanwender und dem Probanden begründet. Der Testanwender kommt hauptsächlich vor der Testung, bei der Werbung der Probanden und dem Verteilen des Fragebogens in Kontakt mit den Probanden. Wird ein Interview als Instrument eingesetzt, so fordert die Gewährleistung der Objektivität besonderen versuchsmethodischen Aufwand: Alle Interviewer müssen über die gleiche Fachkenntnis und Aufmerksamkeit in allen Interviews verfügen.

Bei der Beurteilung der *Validität* kommt es auf das vermessene Konstrukt an. Insbesondere das Konstrukt Wissen kann in Form einer Leistungsbewertung valide erfasst werden. Bei Befragungen kann das Konstrukt des Verhaltens jedoch

nur als Selbstbericht über hypothetisches oder vergangenes Verhalten erfasst werden. Wird hypothetisches Verhalten abgefragt, können sich, insbesondere in Abhängigkeitssituationen wie einem Angestelltenverhältnis, Verzerrungen durch soziale Erwünschtheit ergeben [Edw58; DB16]. Auch sind Fehleinschätzungen der Befragten von möglichen Kosten von Entscheidungen eine Quelle für Verzerrungen der Befragungsergebnisse [Hen09]. Wird vergangenes Verhalten abgefragt, kann dieses nur verzerrt durch die Wahrnehmungskompetenz des Probanden wiedergegeben werden. Dieser Effekt ist als „*Recall Bias*" bekannt [Has05]. Jemand, der die Auswirkungen einer Attacke spürt, etwa abgehende Beträge vom eigenen Konto, kann anhand dessen wahrscheinlich nicht angeben, ob dieses Ereignis auf einen erfolgreichen Phishing-Angriff zurückzuführen ist oder ob er einer anderen Form von Betrug, z. B. einem Identitätsdiebstahl, zum Opfer gefallen ist [FJ07]. Auch kann Scham über eine gelungene Täuschung oder ein Fehlverhalten die Antworten eines Probanden beeinflussen [Kel+12]. Damit ist die *Validität* der Messung dieses Konstrukts als abhängige Variable infrage zu stellen. Auch die Anwender dieser Methode empfehlen die Begleitung von Befragungen durch Experimente [KK06].

Die *Reliabilität* der Messung kann, im Verhältnis zu anderen Methoden, einfach kontrolliert werden. Da mehrere Items pro Konstrukt verwendet werden, können diese auf Konsistenz [DB16] getestet werden. Die Messung mittels eines Fragebogens ist, anders als bei einem Experiment (vgl. Abschnitt 2.2.4), reizunabhängig. Damit kann sich die Messung auch bei erneuter Messung stabil verhalten.

Befragungen werden zumeist eingesetzt, um die Korrelation zwischen verschiedenen Konstrukten aufzuzeigen. Items, die ein bestimmtes Konstrukt adressieren, lassen sich durch ihre textuelle Form mit wenig Aufwand im akademischen Diskurs kommunizieren. Konkret kommt es bei der Bewertung der *Normierung* auf die Implementierung an.

Insbesondere sind fragebogenbasierte Studien *ökonomisch*. Fragebögen können beliebig repliziert und verteilt werden, die Auswertung kann automatisiert werden. Interviewbasierte Studien sind durch den Einsatz eines Interviewers nicht ökonomisch umzusetzen.

Befragungen sind im Rahmen der Forschung sehr *nützlich*. Mit ihnen lassen sich Zusammenhänge einer enormen Bandbreite von Konstrukten oft sehr ökonomisch nachweisen. Der praktische Nutzen dieser Methode zur Bewertung von IT-Sicherheitsbewusstsein ist jedoch aufgrund der mangelnden Validität eingeschränkt.

2.2.3 Laborexperimente

Laborexperimente erlauben die isolierte Beobachtung der Probanden beim Lösen einer gestellten Aufgabe. Dabei wird eine spezifische Umgebung gestaltet, in der ein Proband eine gestellte Aufgabe lösen muss.

Kumaraguru u. a. versetzen 30 Versuchspersonen in die Rolle des Mitarbeiters einer fiktiven Firma. Als solcher muss die Versuchsperson auch die E-Mails dieses Mitarbeiters bearbeiten. Als Instrument kommt hier der Versuchsleiter zum Einsatz, der die Versuchspersonen während des Versuchs beobachtet. Im Lauf des Versuchs kommen E-Mails zum Einsatz, die der Proband als Phishing oder legitime E-Mail erkennen soll. Ob der Proband diese Erkennung korrekt vorgenommen hat, entscheidet der Versuchsleiter. Als Metrik kommt die Erkennungsrate zum Einsatz. [Kum+07]

Parsons u. a. können in einem Laborexperiment zeigen, dass die Tatsache, ob Probanden um das Ziel der Studie wissen, einen Einfluss auf ihre Klassifikation von Phishing-E-Mails hat. Es gelingt ihnen zu demonstrieren, dass Probanden, die über das Ziel des Experiments informiert werden, E-Mails deutlich genauer klassifizieren und sich für den Vorgang mehr Zeit lassen als die Probanden, die nicht über das Ziel der Studie in Kenntnis gesetzt werden. [Par+13b]

Anandpara u. a. demonstrieren, dass eine Intervention mit Fokus auf Phishing dazu führt, dass E-Mails häufiger als Phishing klassifiziert werden (Falsch-Positiv). Im gleichen Versuch demonstrieren sie, dass Klassifikationsaufgaben im Labor die Kompetenz der Probanden, Phishing-E-Mails zu erkennen, nicht valide erfassen. [Ana+07]

Die *Objektivität* beim Einsatz dieser Methode lässt sich nur mit hohem versuchsmethodischem Aufwand gewährleisten. Verglichen mit interviewbasierten Studien ist dieser noch höher, da die Beobachtungen des Testanwenders das Ergebnis der Datenerfassung darstellen.

Die gewonnenen Erkenntnisse sind nur bedingt auf den Alltag übertragbar. Ein Proband kann durch die Testsituation sein Verhalten anpassen [FJ07; Par+13b]. In der Humanforschung ist das Prinzip der *Anforderungsmerkmale* (englisch „*Demand Characteristics*“) wohl bekannt [MdBW12]. Der Proband wird durch das Wissen oder seine Wahrnehmung der Studie und ihrer Ziele beeinflusst. Diese Beeinflussung kann zu unterschiedlichen Effekten führen. Das Experiment von Anandpara u. a. demonstriert darüber hinaus, dass Klassifikationsaufgaben im Labor zur Erfassung der Kompetenz eines Probanden, zwischen Phishing-E-Mails und regulären E-Mails zu unterscheiden, ungeeignet zu sein scheinen. Aus diesen Gründen muss angenommen werden, dass die *Validität* der Labormessungen im Bereich der IT-Sicherheitsbewusstseinsmessung nicht gegeben ist.

Die Kontrolle der Situation und aller Einflussfaktoren wirkt sich positiv auf die *Reliabilität* aus. Im Alltag ist ein Proband jedoch mit unterschiedlichen Aufgaben betraut, die eine unterschiedlich hohe kognitive Belastung bedeuten. Diese Belastung beeinflusst die Leistungsfähigkeit bei aufmerksamkeitsbasierten Aufgaben [Lav10]. Auch Einflussfaktoren wie Stress wirken sich auf die Leistungsfähigkeit der Probanden aus [Gai08]. Dies ist auch der Grund, aus dem die Ergebnisse nicht auf den Alltag der Probanden übertragbar sind. Aus diesem Grund ist die *Nützlichkeit* eingeschränkt.

Darüber hinaus muss angemerkt werden, dass Laborexperimente zur Vermessung von IT-Sicherheitsbewusstsein aufgrund des hohen personellen Aufwands nicht *ökonomisch* sind. Es sind auch diese ökonomischen Anforderungen, die eine mögliche *Normierung* hemmen.

2.2.4 Feldexperimente

Im Rahmen eines Feldexperiments wird der Proband nicht in ein Laborumfeld versetzt; er verbleibt während des Tests in seinem gewohnten Umfeld. Diese Methode kommt auch in der Forschung zu Phishing-E-Mails zum Einsatz [Jag+07; JR06; Cap+14].

Die Agentur für Cybersicherheit der Europäischen Union (kurz: ENISA) schlägt als Metrik zur Bewertung für die IT-Sicherheit die Resilienz der Nutzer gegen Cyberangriffe vor, insbesondere das Ausmaß, in dem Nutzer einen Cyberangriff erkennen. Weiter schlägt sie als Instrument ein Quiz oder

einen computerbasierten Test vor. Ein weiterer Vorschlag ist die Erfassung der Nutzeranzahl, die einem simulierten Angriff zum Opfer fallen. Als Messinstrument wird eine Phishing-E-Mail vorgeschlagen. [Theo7]

Das Feldexperiment wurde zuvor bereits von Hasle u. a. eingesetzt, um die Resistenz von Probanden gegen Social Engineering zu messen. Als Instrument werden E-Mails genutzt. Die erste E-Mail gibt vor, zu einer Umfrage einzuladen, die zweite E-Mail hat keinen lesbaren Inhalt, öffnet jedoch einen Dialog, welcher den Nutzer auffordert, sein Passwort einzugeben. Ein Proband wird als aktiv registriert, wenn die E-Mail gelesen wird. Um den Status der E-Mail festzustellen, wird der Verweis auf eine Grafik in der E-Mail hinterlegt. Dieser Verweis ist mit einem eindeutigen Identifikator versehen. Wird die Grafik zur Darstellung der E-Mail vom Webserver abgerufen, kann diese Anfrage aufgezeichnet werden und erlaubt auf diese Weise Rückschlüsse auf den Status der E-Mail. Gemessen wird der Anteil der Nutzer, der die E-Mail öffnet, der auf den Link klickt (oft auch englisch „*click rate*“) sowie der Anteil der Nutzer, der ein Passwort eingibt (oft auch „*input rate*“). [Has+05]

Steyen, Kruger & Drevin manipulieren das E-Mail-Programm aller Studenten einer südafrikanischen Universität, um festhalten zu können, wie viele Nutzer die versendete Phishing-E-Mail öffnen, auf sie antworten, auf den angebotenen Hyperlink klicken oder auf der verknüpften Website ihr Passwort eingeben. [SKD07]

Dodge, Carver & Ferguson nutzen multiple Phishing-E-Mails, um die Auswirkungen jährlicher IT-Sicherheitsbewusstseinsmaßnahmen unter Studierenden einer Militärakademie zu illustrieren. Hierzu wird der Anteil der Nutzer, der auf den in der E-Mail angebotenen Link klickt und der Anteil der Nutzer, der den Empfang der E-Mail meldet (oft auch englisch „*report rate*“), erfasst. Es kann demonstriert werden, dass beide Werte mit der Anzahl der Studienjahre korrelieren. Die relative Anzahl der Meldungen der Studenten korreliert positiv mit der Anzahl der Studienjahre. Die relative Anzahl der Studenten, der auf den Link klickt, korreliert negativ mit der Anzahl der Studienjahre. [DCF07]

In einer Studie unter 311 Mitarbeitern eines Unternehmens wird die Wirksamkeit einer Interventionsmethode in Form eines Comics durch den Einsatz von Phishing-E-Mails untersucht [Kum+08]. Auch wenn die Wirksamkeit

dieser Intervention nicht belegt werden kann, können die Testanwender die Beobachtung machen, dass die Probanden das Schulungsmaterial untereinander teilten.

Ein weiteres Feldexperiment im universitären Umfeld mit Fokus auf den Täuschungsvorgang in Phishing-E-Mails nutzt alle Reaktionen eines Empfängers, die darauf schließen lassen, dass der Nutzer den Täuschungsversuch erkannt hat. Aus diesem Grund wird die *report rate* mit hohem versuchsmethodischen Aufwand erfasst. Dazu gehören Meldungen an den technischen Support, genauso wie Antworten auf die E-Mail mit einem Ausdruck des Unglaubens oder Misstrauens. Weitere Aktionen der Probanden wie Interaktionen mit der E-Mail werden nicht aufgezeichnet. [Wri+10]

Wolf, Haworth & Pietron quantifizieren den Effekt einer IT-Sicherheitsbewusstseinsmaßnahme, in der Kriterien für „gute“ Passwörter vermittelt werden. Der Effekt wird anhand der Komplexität der nutzervergebenen Passwörter gemessen. Sie implementieren eine Eingruppen-Pretest-Posttest-Studie [SCC02]. Ein positiver Effekt der Intervention kann demonstriert werden. Darüber hinaus kann demonstriert werden, dass weitere Interventionen der gleichen Gruppe keine weitere Verbesserung erbringt. [WHP11] Auch Eminağaoğlu, Uçar & Eren setzen Passwörter zur Bewertung von IT-Sicherheitsbewusstsein ein [EUE09].

Abseits des akademischen Diskurses empfiehlt Gréaux zur Bewertung der Gesamtheit der Nutzer die Zeitspanne von dem Verschicken einer Phishing-E-Mail bis zum ersten Öffnen des Links sowie der ersten Meldung der Phishing-E-Mail über eine im E-Mail-Client angebotene Schaltfläche [Gré15].

Wie stark der Testanwender in den Test involviert ist, hängt von der eingesetzten Variante des Experiments ab. Die Beobachtung des Probanden bei der Arbeit durch den Testanwender ist zwar theoretisch möglich [HB14], jedoch mit immensen Kosten verbunden. Darüber hinaus ist bereits aus der psychologischen Forschung bekannt, dass Probanden ihr Verhalten in Beobachtungssituation, verändern können. Man spricht von dem *Hawthorne Effekt* [MWE14]. Außerdem ergeben sich Fragestellungen nach datenschutzrechtlichen Implikationen und der Vertraulichkeit betrieblicher Interna durch den weitreichenden Einblick, der dem Testanwender in die Arbeitsumgebung gewährt wird. Aus diesen Gründen ist die manuelle Beobachtung der Probanden im Tagesbetrieb nicht anwendbar.

Bei einem voll automatisierten Experiment tritt der Testanwender, ähnlich dem Vorgehen bei einer fragebogenbasierten Studie, nur sehr eingeschränkt in Kontakt mit den Probanden. Gerade wenn Phishing-E-Mails als Messinstrument zum Einsatz kommen, wird teils empfohlen, auf Täuschung der Probanden zurückzugreifen [FJ07], oder es wird vor Abschluss der Testung von der Kommunikation mit den Probanden abgesehen [JR06]. Damit liegt der Methode Potenzial zur *Objektivität* inne.

Die Messung erfasst das tatsächliche Verhalten der Probanden, was für die *Validität* der Messungen mit dieser Methode spricht. Dem ist entgegenzuhalten, dass das sicherheitsrelevante Verhalten der Probanden vollumfänglich erfasst werden muss. Die hier dargestellten Feldexperimente zu Phishing geben keinen Aufschluss darüber, ob der Proband die Möglichkeit hat den Vorfall zu melden, falls die Interaktion mit der Phishing-E-Mail im Vordergrund steht [DCF07; SKD07; Has+05]. Steht die Erkennung der Phishing-E-Mail im Vordergrund, werden die Interaktionen mit der Phishing-E-Mail nicht vollständig erfasst [Wri+10]. Weiter spricht gegen die Validität der Erfassung von IT-sicherheitsrelevantem Verhalten, dass nur das Verhalten in Bezug auf Phishing-E-Mails und Passwörter in Feldexperimenten untersucht wurde. Es ist infrage zu stellen, ob das untersuchte Verhalten als tatsächlicher Repräsentant für IT-Sicherheitsbewusstsein tragbar ist.

Ob diese Feldexperimente *reliabel* sind, ist hier nicht bewertbar. Da häufig nur eine kleine Anzahl an Phishing-E-Mails verschickt wird und diese nicht in Beziehung zueinander gesetzt werden, fehlt die zur Bewertung erforderliche Datengrundlage. Die Studie zum Verhalten von Nutzern bezüglich der Wahl von Passwörtern wiederholte die Messung täglich über einen Zeitraum von 14 Tagen und demonstrierte, dass sich nur an einem Tag nach der Intervention eine Veränderung nachweisen lässt [WHP11]. Das ist nicht verwunderlich, da die Passwortwahl durch Nutzer nur in größeren zeitlichen Intervallen und selten anlasslos stattfindet [SB14; FH07]. Bisherige akademische Dokumentationen von Feldexperimenten liefern keinen Aufschluss darüber, ob sich das Verhalten der Nutzer bezüglich des IT-Sicherheitsbewusstseins mit dieser Methode *reliabel* messen lässt.

Feldexeperimente geben jede gemessene Variable einzeln an und es fehlt an *Skalierung*. Diese Methode ist bei gegebener Automatisierung *ökonomischer* als

Laborexperimente, aber nicht so ökonomisch wie Studien. Durch die Tatsache, dass das konkrete Verhalten innerhalb einer für den Probanden alltäglichen Situation vermessen wird, ist die Methode der Feldexperimente zur Bewertung der Effekte von IT-Sicherheitsbewusstseinsmaßnahmen auf das Verhalten der Nutzer durch die hohe Validität besonders *nützlich*.

Zwar suggerieren die erfassten Werte *click rate*, *input rate* und *report rate* durch ihr wiederkehrendes Auftreten eine *Normierung*, es ist jedoch davon auszugehen, dass diese Werte vom Reizgeber, also z. B. der konkreten Phishing-E-Mail, abhängig sind. Die Anzahl der unterschiedlichen Werte zeigen, dass diese Methode bisher keine *Skalierung* anbietet. Durch die Validität der Erfassung des Verhaltens und der Implikationen dieser für die Sicherheit einer Organisation ist eine derartige Messung besonders *nützlich*.

2.2.5 Sicherheitsmetriken

Wie bereits in Abschnitt 2.1.1 dargestellt, ist die Erhöhung der IT-Sicherheit die Motivation von Interventionen und der Forschung in diesem Bereich. Oft wird mit einer Steigerung des IT-Sicherheitsbewusstseins eine Reduktion von IT-sicherheitsrelevanten Vorfälle assoziiert [WH03; Bun20c; The10]. Damit einhergehend soll der Anteil der erfassten IT-sicherheitsrelevanten Vorfällen, der durch Nutzer gemeldet wurde, steigen [WH03; Bun20c; The10]. Vorab kann direkt festgestellt werden, dass es sich bei dieser Methode durch die direkte Verknüpfung mit der Motivation um eine *nützliche* Methode handelt.

Auch die ENISA gibt an, dass typische Metriken die Anzahl, die verursachten Kosten oder die Ausfallzeiten von nutzerbedingten IT-Sicherheitsvorfällen sind. Es wird angegeben, dass manchen Unternehmen diese Zahl mit der Gesamtanzahl aller IT-Sicherheitsvorfälle relativieren. [The10]

Im Vergleich zu den anderen Messmethoden, in denen wohlbekannte Fälle konstruiert werden, liegt dieser Messmethode eine Datenbasis zugrunde, deren wahre Klassifikation unbekannt ist. Jede Meldung der Nutzer muss auf ihre Relevanz für die IT-Sicherheit geprüft werden. Diese Klassifikation übersteigt die Kapazitäten eines einzelnen Testanwenders. Mehrere Testanwender müssten die Klassifikation durchführen und notwendige Recherchen anstellen. Damit ist eine *objektive* Umsetzung dieser Methode nur mit enormem versuchsmethodischen

Aufwand zu realisieren und steht damit der *Testökonomie* dieser Methode direkt entgegen.

Die Güte der Erfassung aller IT-sicherheitsrelevanten Vorfälle hängt maßgeblich von der Güte der unterstützenden IT-Systeme ab. Die absolute Anzahl der IT-sicherheitsrelevanten Vorfälle ist hängt ihrerseits von der aktuellen Bedrohungslage ab. Es ist an dieser Stelle nicht zu bewerten, wie stark diese Einflüsse auf das entsprechende Maß einwirken und ob sich die Auswirkungen der IT-sicherheitsbewusstseinsfördernden Maßnahme quantifizieren lassen. Die *Validität* dieser Methode kann hier nicht bewertet werden. Unter diesen Umständen erübrigt sich eine weitere Betrachtung.

2.2.6 Hybride Ansätze

Hybride Ansätze können unter Umständen die Stärken zweier Messmethoden kombinieren. So können die mit Fragebögen erfassbaren Konzepte mit der Verhaltenserfassung der Feldexperimente kombiniert werden. Auf diese Weise ist es möglich, tatsächliches Verhalten eines Probanden mit individuellen Einflussfaktoren zu korrelieren.

Halevi, Memon & Nov kombinieren Fragebögen zur Erfassung von Persönlichkeitsmerkmalen, den *Big Five* der Probanden, mit einem Phishing-Feldtest, bei dem dem Empfänger ein Preis in Aussicht gestellt wurde. Für Frauen konnte eine Korrelation zwischen der Empfänglichkeit für Phishing und Neurotizismus festgestellt werden. Bei Männern korrelierte keines der fünf Persönlichkeitsmerkmale mit der Empfänglichkeit für Phishing. [HMN15]

Workman, Bommer & Straub nutzen Fragebögen für selbstberichtetes Verhalten und psychologische Konzepte wie die Bedrohungseinschätzung und verschiedenen Bewältigungsstrategiekonzepten. Darüber hinaus extrahieren die Autoren Nutzerverhalten aus Log-Dateien. Das extrahierte Verhalten ist das regelmäßige Wechseln des Passworts, die Sicherung von unternehmensrelevanten Daten und die manuelle Ausführung eines Virusscans. Es konnte eine starke Korrelation zwischen der Kontrollüberzeugung (extern oder intern) und dem selbstberichteten Verhalten gezeigt werden. Diese Korrelation war nicht zu dem tatsächlichen Verhalten der Probanden nachweisbar. [WBS08]

Auch wenn mit der Kombination von Befragung und Feldversuch der Mangel an *Validität* der Befragung kompensiert werden kann, bleibt die Kritik an Feldversuchen, wie in Abschnitt 2.2.4 auf Seite 25 vorgebracht, bestehen.

2.3 Fazit

Damit ist die erste Forschungsfrage nach den Stärken und Schwächen bisher eingesetzter Verfahren zur Messung von IT-Sicherheit (vgl. Abschnitt 1.3) beantwortet. Tabelle 1 gibt einen Überblick.

Tabelle 1: *Bewertung der zu Messung von IT-Sicherheitsbewusstsein eingesetzten Methoden bezüglich etablierter Gütekriterien psychologischer Tests. Die Bewertung ist wie folgt codiert – „+“: Methode zeigt Stärken in diesem Kriterium, „○“: Methode zeigt in diesem Kriterium keine Stärke oder Schwäche, „–“: Methode zeigt Schwächen in diesem Kriterium; wurde keine direkte Aussage abgeleitet, ist kein Zeichen eingetragen.*

	Objektivität	Veladität	Reliabilität	Skalierbarkeit	Normierung	Testökonomie	Nützlichkeit
Befragungen (Fragebogen)	+	–	+	+	○	+	○
Befragungen (Interview)	○	–	+	+	○	–	○
Laborexperimente	○	○	+	○	–	–	○
Feldexperimente	+	+		–	○	○	+
Sicherheitsmetriken	–	–				–	+
Hybride Ansätze	+	+		–	○	○	+

Der Einsatz von Sicherheitsmetriken zur Quantifizierung von IT-Sicherheitsbewusstsein der Nutzer präsentiert sich besonders nützlich. Dieses Vorgehen ist jedoch kaum valide zu umzusetzen. Feldexperimente sind die einzige Methode, die Nutzerverhalten als zentrales Konzept des IT-Sicherheitsbewusstseins (siehe Abbildung 2 auf Seite 15) direkt und damit valide erfasst. Damit bieten sie das größte Potenzial für weitere Entwicklung. Leider werden Feldexperimente fast ausschließlich mit Phishing-E-Mails umgesetzt. Dies liegt vermutlich in der Tat-

sache begründet, dass Phishing-E-Mails sich exakt an die Probanden adressieren lassen [Cro+13].

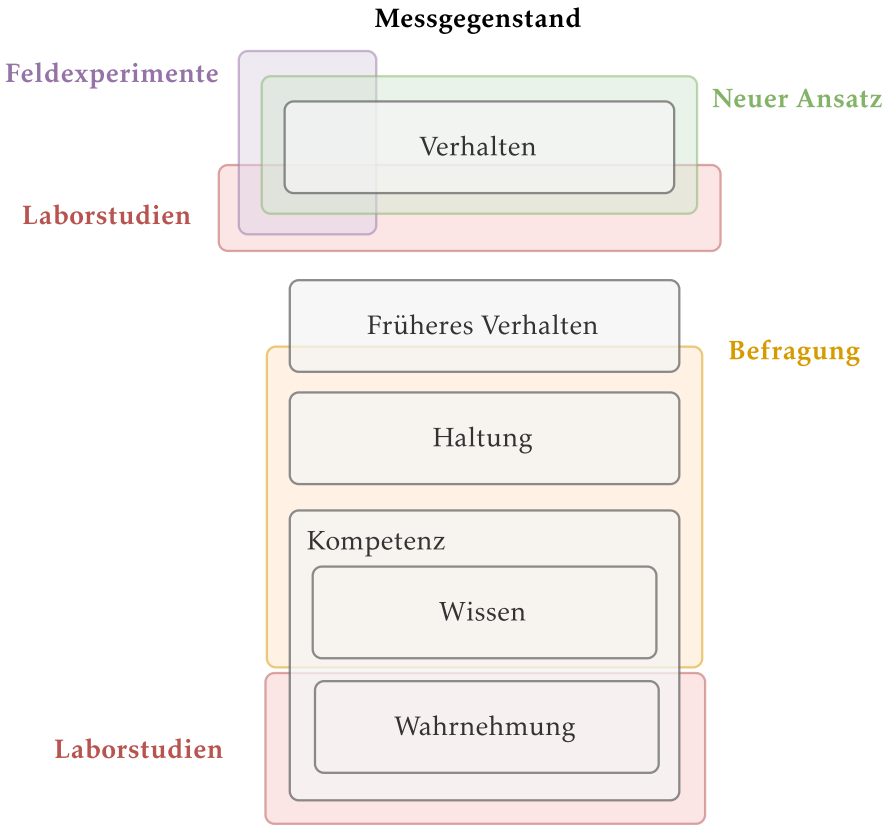


Abbildung 3: *Abdeckung der verschiedenen Messgegenstände durch Methoden der IT-Sicherheitsbewusstseinsforschung.*

Abbildung 3 zeigt die Abdeckung verschiedener Messmethoden bezüglich der Konzepte des IT-Sicherheitsbewusstseins. Befragungen können zur Messung der Haltung und des Wissens des Nutzers eingesetzt werden. Das gesamte Konzept der Kompetenz kann nicht erfasst werden. Eine Befragung lässt nur wenig Rückschluss auf die Fähigkeit zur Wahrnehmung des Nutzers zu. Aber genau diese Fähigkeit des Nutzers wird vorausgesetzt, wenn früheres Verhalten abgefragt wird. Laborstudien erlauben die Vermessung der Wahrnehmungskompetenz der Probanden. Das gesamte Verhalten eines Nutzers ist aber mit dieser Methode nicht valide zu erfassen [Ana+07]. Feldexperimente können das Verhalten der Nutzer zwar valide erfassen, wurden aber bisher nur auf einem Teilbereich des IT-sicherheitsrelevanten Verhalten erprobt. Die Stärken

der Methoden lassen sich mit einem erweiterten, feldexperimentellen Ansatz vereinen. Hier liegt die Chance in einer validen und objektiven Messung von Nutzerverhalten, die sonst in keiner Methode gegeben ist.

Um die Messung von IT-Sicherheitsbewusstsein anhand von IT-sicherheitsrelevantem Verhalten zu ermöglichen, muss ein breiteres Spektrum an Verhalten erfasst werden. Dies wird mithilfe einer Generalisierung der Phishing-Tests zu *artefaktbasierter Messung* erreicht. Dazu wird untersucht, welche Reizgeber neben Phishing-E-Mails noch in Betracht kommen, um Nutzerreaktionen zu provozieren. Die Herleitung ist in Kapitel 3 beschrieben. Die artefaktbasierte Messung kann mit Fragebögen kombiniert werden, um den Einfluss verschiedener psychologischer Konzepte auf das tatsächliche Verhalten zu messen.

Zur Demonstration der Validität der *artefaktbasierten Messung* kann eine Intervention eingesetzt werden. Mit fast allen in der Forschung fest etablierten Methoden lässt sich ein positiver Effekt der Intervention illustrieren. Eine Ausnahme ist das durch Anandpara u. a. durchgeführte Laborexperiment (vgl. Abschnitt 2.2.3). Es ist nicht in der Lage, die Fähigkeit zur Erkennung von Phishing valide zu messen. Khan u. a. stellten fest, dass sich Gruppendiskussionen und Frontalunterricht als besonders effektiv präsentieren [Kha+11]. Eine Umfrage unter 67 europäischen Unternehmen ergab, dass ein Seminar die am effektivsten geglaubte Methode ist, um das IT-Sicherheitsbewusstsein zu steigern [Theo7].

Können mehrere Reizgeber zur Messung des IT-Sicherheitsbewusstseins eingesetzt werden, lässt das eine Bewertung der Reliabilität der Messung zu. Eine aktuelle Meta-Studie unter etwa 70 Nutzern über den Zeitraum von fünf Jahren zeigt ein durchwachsenes Bild bezüglich der *click rate* von insgesamt 15 unterschiedlichen Phishing-E-Mails [Gre+18]. Dies illustriert die Abhängigkeit des Messergebnisses von dem Reizgeber und unterstreicht den Bedarf nach einer Untersuchung der Reliabilität für Feldversuche.

Damit sich die neue Methode ökonomisch implementieren lässt, muss sich diese umfassend automatisieren lassen. Der Entwurf eines Werkzeugs wird in Kapitel 5 beschreiben. Zur Skalierung der Messmethode muss *ein* Testwert die Interpretation der Messung erlauben. Dieses wird in Abschnitt 3.4 hergeleitet. Damit werden die Schwächen bisheriger Feldversuche bis auf die Normierung kompensiert.

3 Artefaktbasierte IT-Sicherheitsbewusstseinsmessung

Im Rahmen dieses Kapitels wird das Prinzip der artefaktbasierten IT-Sicherheitsbewusstseinsmessung hergeleitet. Diese Darstellung ist fundamental für die Beantwortung der zweiten Forschungsfrage nach einem neuen Ansatz, der die Stärken bisheriger Ansätze vereint (vgl. Abschnitt 1.3). Zunächst wird das *Situationsmodell* dargestellt, das der Nutzung eines Bildschirmarbeitsplatzes zugrunde liegt. Darauf folgend wird das Konzept des *Artefakts* als künstliches Element in einer Situation definiert. Anschließend werden die *Handlungsoptionen* des Nutzers in Bezug auf das Artefakt vorgestellt, bevor abschließend *Maße* zur Bewertung der ergriffenen Handlungsoption hergeleitet werden.

3.1 Situationsmodell

Bedient ein Nutzer seinen Bildschirmarbeitsplatz, durchläuft der Nutzer eine *Situation*. Dabei verfolgt der Nutzer ein oder mehrere Zielstellungen [Bre92]. Um diese zu erreichen, manipuliert der Nutzer das System durch Interaktion mit den *Elementen* der Situation [End88]. Elemente sind *wahrnehmbare* Objekte in der jeweiligen Situation. Elemente können eine Interaktionsoption anbieten, z. B. eine Schaltfläche, die per Mausklick betätigt werden kann. Dabei ist der Erfolg von einer Serie aus Entscheidungen abhängig und jede Interaktion kann die Situation und damit ihre Elemente ändern [Bre92]. Das Verhalten von Nutzern wird, wie in Kapitel 2 dargestellt, durch äußere Einflüsse, die *Testumgebung*, beeinflusst, in der die Entscheidungen getroffen werden. Im Kontrast zu den Zielstellungen aus der Verhaltensforschung, in den dynamischen Systemen der klassischen Entscheidungstheorie [Bre92], folgen die getroffenen Entscheidungen nicht primär dem Ziel, die Sicherheit des IT-Systems und damit mittelbar der

gesamten IT-Infrastruktur zu erhalten oder zu verbessern. Vielmehr wird der Bildschirmarbeitsplatz dazu genutzt, eine unabhängige Zielstellung zu verfolgen.

Ein Teil der Elemente einer solchen Situation hat einen Bezug zur IT-Sicherheit des Systems. Interaktionen mit diesen Elementen können damit einen Einfluss auf die Sicherheit der IT-Infrastruktur haben. Durch diese Elemente lässt sich eine Subsituation definieren. Während der Nutzer sein Primärziel verfolgt, muss er auch mit Elementen mit Sicherheitsbezug interagieren. Abbildung 4 illustriert dieses Konzept.

Das Verhalten des Nutzers in dieser Subsituation wird als Ausdruck von IT-Sicherheitsbewusstsein verstanden. Zur kontrollierten Erfassung des Verhaltens kann nun ein Element *künstlich* in diese Situation eingegeben werden. Aufgrund dieser künstlichen Eingabe wird zur Referenzierung dieses Elements der Begriff *Artefakt* verwendet. In Abgrenzung zu dem Artefaktbegriff werden die Elemente mit Sicherheitsbezug, die keine Artefakte sind, *natürliche Elemente* genannt.

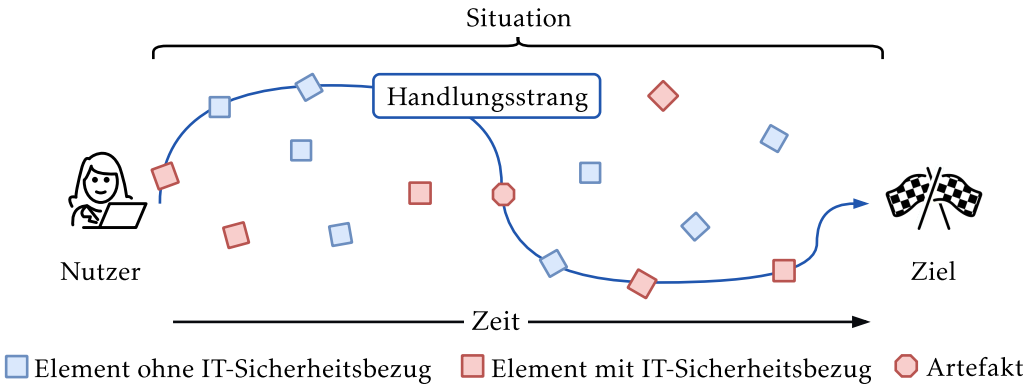


Abbildung 4: *Visualisierung der Situation eines Nutzers bei der Bildschirmarbeit. Elemente der Situation ohne IT-Sicherheitsbezug sind blau dargestellt. Elemente mit IT-Sicherheitsbezug sind rot dargestellt. Natürliche Elemente sind als Quadrate dargestellt, das Artefakt als Oktagon. Der Nutzer interagiert mit den Elementen in chronologischer Abfolge. Es ergibt sich ein „Handlungsstrang". Sollte der Nutzer auch mit dem in die Situation eingebrachten Artefakt interagieren, so liegt auch dieses auf dem Handlungsstrang.*

Theoretisch ist die Messung mit beiden Typen von Elementen möglich. Die Messung des IT-Sicherheitsbewusstseins mit natürlichen Elementen gestaltet sich jedoch unter praktischen Gesichtspunkten problematisch.

Innerhalb einer administrierten IT-Infrastruktur sind natürliche Elemente so zu konfigurieren, dass eine Interaktion mit diesen Elementen die Sicherheit der IT-Infrastruktur erhält. Beispielsweise lassen sich Anforderungen an die Komplexität von Passwörtern programmatisch durchsetzen. Die Durchsetzung dieser Anforderung ist dem Transfer der Verantwortlichkeit auf den Nutzer vorzuziehen. Die Erfassung und Bewertung des Verhaltens des Nutzers in Bezug auf natürliche Elemente ist ebenso zweifelhaft. Zum einen liegen diese Elemente unter der Gestaltungskontrolle der IT-Infrastruktur-Betreiber, zum anderen ist eine nicht-sachgemäße oder unerwünschte Nutzung auch auf fehlerhafte Gestaltung zurückführbar. Aus diesen Gründen ist die Messung von IT-Sicherheitsbewusstsein auf Basis von Artefakten der von natürlichen Elementen vorzuziehen.

Die Präsentation eines Artefakts, die Aufnahme des korrespondierenden Nutzerverhaltens und dessen Interpretation bildet *einen Test*. Wie in Abschnitt 2.3 bereits dargestellt, sollen Nutzer im Rahmen einer Messung von IT-Sicherheitsbewusstsein mehrere Tests durchlaufen.

3.2 Artefakte

Im Rahmen dieses Unterkapitels wird eruiert, welche Artefakte zur Messung von IT-Sicherheitsbewusstsein eingesetzt werden können. Zu diesem Zweck werden zunächst allgemeine Anforderungen an Artefakte dokumentiert. Anschließend werden diese exemplarisch untersucht, um die Anforderungen an ein Werkzeug zur Messung von IT-Sicherheitsbewusstsein festlegen zu können.

Um als Artefakt in einem Test zur IT-Sicherheitsbewusstseinsmessung verwendet werden zu können, muss ein Artefakt zwei grundlegende Eigenschaften erfüllen. Erstens: *Ein Artefakt ist ein Situationselement mit IT-Sicherheitsbezug*. Dem Nutzer muss bewusst sein können, dass diesem Element ein IT-Sicherheitsbezug inhärent ist. Sollte dem Element kein IT-Sicherheitsbezug inhärent sein, dann ist die Nutzerreaktion nicht auf das IT-Sicherheitsbewusstsein zurückzuführen und das Element ist damit zur Testung ungeeignet.

Beeinträchtigt ein Artefakt die Verfügbarkeit einer Ressource, die vom Nutzer zur Erreichung seines Primärziels benötigt wird, so ist die Zielerreichung des

Nutzers nicht mehr gegeben. Im Fall der Nicht-Verfügbarkeit dieser Ressource wird der Nutzer zur Ausnahmehandlung gezwungen. Ein Rückschluss auf das IT-Sicherheitsbewusstsein als Motivation für die Handlung des Nutzers ist nicht mehr möglich. So lautet die zweite Eigenschaft: *Ein Artefakt darf die Nicht-Verfügbarkeit einer kritischen Ressource nicht verursachen.*

Angreifer sind gezwungen ihr Vorgehen konstant an eine fortschreitende Entwicklung der technischen IT-Sicherheitslösungen anzupassen (vgl. Kapitel 1). Darüber hinaus können neue technische Entwicklungen auch neue Angriffsmöglichkeiten bieten. Da sich das Angreiferverhalten damit wie die technische Entwicklung allgemein in einem konstanten Wandel befindet, kann keine unveränderliche, erschöpfende Liste aller möglichen zur Testung geeigneten Artefakte existieren. Um jedoch möglichst präzise Anforderungen an das Werkzeug ermitteln zu können, müssen konkrete Beispiele herangezogen werden.

Zunächst kommen situative Elemente infrage, mit denen bereits zuvor IT-Sicherheitsbewusstsein gemessen wurde. Danach werden Angriffsschemata analysiert, die eine Nutzerinteraktion implizieren. Zuletzt werden Beispiele zusammengetragen und ausgewertet, in denen der Nutzer einem Angriff ausgesetzt wurde.

3.2.1 Artefakte aus der Forschung zu IT-Sicherheitsbewusstsein

Wie bereits in Kapitel 2 dargestellt, wird wiederkehrend der Umgang mit *Passwörtern* [WHP11; EUE09; Sta+05; WBS08] als auch *Phishing-E-Mails* [Cap+14; Wri+10; DCF07] mit IT-Sicherheitsbewusstsein assoziiert. Aber auch der Umgang mit Anti-Virus-Software und Back-up-Software wird mit IT-Sicherheitsbewusstsein in Verbindung gesetzt [WBS08].

Diese Elemente eignen sich jedoch nicht uneingeschränkt zur artefaktbasierten IT-Sicherheitsbewusstseinsmessung. Der Umgang mit natürlichen Elementen sollte, wie in Abschnitt 3.1 dargestellt, nicht zur Messung herangezogen werden. Damit entfallen die Anti-Virus- und die Back-up-Software sowie Passwörter als Reizgeber für die Probanden. Jedoch kann ein Artefakt in gezielter Wechselwirkung mit natürlichen Elementen stehen. Artefakte können Warnungen in natürlichen Elementen auslösen (vgl. Abschnitt 3.2.4 auf Seite 38).

Diese Wechselwechselwirkung ist insbesondere bei Phishing zu finden. Mit Phishing wird ein Angriff auf die Vertraulichkeit schützenswerter Informationen bezeichnet [Cla07; Gar+07; KK05; MKK08; JM06]. Zu derartigen Informationen sind insbesondere das Passwort oder andere Authentifikationsmittel zu zählen [Cla07; KK05; Gar+07].

Für die Durchführung eines Phishing-Angriffs muss der Angreifer mit dem potenziellen Opfer in Kontakt treten. Wie bereits in Kapitel 2 dargestellt, kommt dabei zumeist eine E-Mail als Medium zum Einsatz. Aber auch andere Kommunikationsmedien werden in Phishing-Angriffen verwendet. Es werden Fälle berichtet, in denen die in Computerspielen angebotene Chatfunktion als Kommunikationsmedium für einen Phishing-Angriff genutzt wird [Zor10]. Auch über soziale Medien wie Facebook werden Phishing-Angriffe unternommen [Art09]. Phishing über das Telefon ist auch Teil des akademischen Diskurses [CLC17].

Dabei wird aber nicht nur das Passwort abgefragt. E-Mails eigenen sich ebenso zur Verteilung von Schadsoftware [Mil07, 3.6 Specialized Malware]. Diese wird zum Teil über präparierte Webseiten verteilt [Eck13, Drive-by-Download, S. 164] oder als Download angeboten [Eck13, Trojanisches Pferd, S. 74].

Zur Steigerung der Erfolgsaussichten kann ein Angreifer den Kontext des Adressaten aufgreifen [Jag+07; Jak05]. Dabei kann der Angreifer die Identität des Adressaten verwenden und ihn direkt ansprechen, um zu suggerieren, dass der Absender mit dem Adressaten vertraut ist [Jak05; DKK14] oder die E-Mail eines Dienstes nachahmen, bei dem der Adressat Kunde ist [Cap+14]. Bei einer umfänglich personalisierten Phishing-E-Mail spricht man von *Spear Phishing* [Cap+14; HMN15; BGL17]. Schadsoftware kann sich auf diese Weise auch automatisiert verbreiten, indem sie auf empfangene E-Mails antwortet [CI17].

3.2.2 Artefakte in konzeptionell bedingten Nutzerschnittstellen von Angriffen

Es existieren ganze Klassen von Angriffen, die konzeptionell eine natürliche Schnittstelle zum Nutzer innehaben. Dies ist bei den bereits im vorigen

Abschnitt besprochenen Phishing-Angriffen der Fall. Oft sind Angriffstypen oder Schadsoftwareklassen nach ihrer Schnittstelle zum Nutzer benannt.

Ein prominentes Beispiel sind *Denial-of-Service-Angriffe* [Eck13, S. 121]. Hier wird dem Nutzer eine Ressource, meist durch das Erzeugen von Überlastsituationen, entzogen. Eng verwandt mit diesem Angriff ist das „*Defacing*“ von Webseiten. Bei diesem Angriff wird die Webpräsenz einer Institution so verändert, dass Besuchern der Webpräsenz ein anderer, oft diffamierender Inhalt präsentiert wird [RvdH17].

Schadsoftwareklassen sind häufig nach ihrer Schnittstelle zum Nutzer benannt. So existiert eine Schadsoftwareklasse, die durch gezielte Präsentation von Werbung Einnahmen generiert. Dies kann beispielsweise durch Pop-ups oder die Änderung der Startseite des Browsers geschehen. Schadsoftware dieser Art ist als *Adware* bekannt [Ori16, S. 237; FG06]. Soll der Nutzer hingegen mit bedrohlich wirkender Rhetorik dazu gebracht werden, eine bestimmte Aktion auszuführen, spricht man von *Scareware* [Ori16, S. 237].

Es existieren auch Vorgehensweisen, die eine direkte Monetarisierung erlauben. *Ransomware* [Ori16, S. 238] kann Nutzerdaten verschlüsseln. Das Original wird anschließend gelöscht. Für deren Wiederherstellung wird ein Lösegeld gefordert. Dabei ist die Aktivität dieser Schadsoftware nicht nur durch die Lösegeldforderung zu erkennen. Viele Exemplare sind an der Gestalt der Kopien der Dateien, den Dateinamen oder deren Endungen zu erkennen [Int14; Abr16].

Mit dem Aufkommen von Kryptowährung entstand auch eine neue Schadsoftwareklasse, die *Miner* [Hua+14]. Diese nutzen entweder die Rechenkraft von vielen kompromittierten Computern oder den Besuchern einer kompromittierten Webseite durch eine Implementierung in JavaScript. Nutzer können die Auswirkungen der erhöhten Leistungsaufnahme, z. B. erhöhte Lüfteraktivität oder erhöhte Reaktionszeiten auf Nutzereingaben, erfahren.

3.2.3 Artefakte in öffentlich bekannt gewordenen Angriffen

Immer wieder kommt es dazu, dass Angriffe eine derart spektakuläre Gestalt annehmen, dass sie in einschlägigen Nachrichtenplattformen Erwähnung finden. Auch finden sich Benutzerberichte auf Community-Plattformen, in denen Nutzer um Hilfe oder Rat fragen, um zu ermitteln, ob eine gemachte Beobachtung eine

Anomalie darstellt oder ob diese Beobachtung einem verbreiteten Phänomen zuzuordnen ist. Anbieter von Sicherheitssoftware und -dienstleistungen werben mit aufgeklärten oder abgewandten Angriffen. Einige dieser Berichte werden folgend vorgestellt, weil sich damit der Ursprung der betrachteten Artefakte illustrieren lässt, was einen Hinweis auf weitere mögliche Gestalten liefert.

Palo Alto Networks Inc. hat 2014 einen Bericht zu einem Vorfall der Schadsoftware „Wirelurker“ herausgegeben [Xia14]. Der Aufklärungsprozess wurde durch einen Nutzerbericht angestoßen. Dieser Nutzer fand eine unbekannte Applikation auf seinem Mac und seinem iPhone. Die Analyse ergab, dass 467 unterschiedliche Applikationen mit der Schadsoftware infiziert waren und über den App Store eines Drittanbieters heruntergeladen werden konnten. Mac-Applikationen werden typischerweise in Form von Datenträgerabbildern distributiert [Sino6, 11.4.1. Using the `hdiutil` Program]. Alle Datenträgerabbilder der infizierten Software hatten ein spezifisches Hintergrundbild gesetzt (Abbildung 5 zeigt dieses). Es handelt sich dabei um eine Art „Handschrift“ und soll durch den Nutzer wahrgenommen werden.

Abbildung 5: *Typische Präsentation eines mit Wirelurker infizierten Programms im Falle CleanApp [Xia14].*

Im Juni 2016 [Red16] konnten Nutzer miterleben, wie ihre Computer scheinbar selbsttätig Anmeldeversuche bei PayPal (`https://paypal.com`) und anderen Onlinediensten unternahmen. Dieses Verhalten wurde durch Angreifer ausgelöst, die sich illegitim Zugriff zu den Anmeldedaten des Dienstes *TeamViewer* der

Nutzer verschafft hatten. TeamViewer ist eine Softwarelösung für „Fernzugriff & Fernwartung“ [Tea20]. Hier gingen die Angreifer ein *kalkuliertes Risiko* ein, dass der Nutzer die Aktivität vorzeitig wahrnimmt.

Dieses Verhalten ist insbesondere gegenüber einem *Vorgehensfehler*, den man etwa bei orthografischen Fehlern in der Kommunikation der Angreifer vermuten kann, abzugrenzen. Diese werden von Nutzern häufig zur Entscheidung herangezogen, ob es sich bei einer E-Mail um einen Betrugsversuch handelt oder nicht [Jak07; DHC06]. Dabei handelt es sich aber nicht unbedingt um das Unvermögen der Angreifer. Damit eine E-Mail den Nutzer erreicht, muss sie zunächst einen *Spamfilter* passieren. Dabei handelt es sich um eine Softwarelösung, deren Aufgabe es ist, ungebetene und ungewollte E-Mails zu filtern. Diese Spamfilter sind jedoch anfällig gegenüber orthografischen Alternativen [Bra+06, 7.4 Sensitivity to Noise in the Data]. Dies legt die Vermutung nahe, dass Angreifer ihre Texte bewusst derart gestalten, um diese Filter zu überwinden. Herley mutmaßt weiter, dass es sich bei der Wahl der Textgestaltung um einen Selektionsprozess der Angreifer handeln kann [Her12]. Er argumentiert, dass die nachfolgenden Schritte des Angriffs weitere Kosten auf Seiten des Angreifers erzeugen. Um diese Kosten zu minimieren, muss er nach der ersten Kontaktaufnahme die Personen aus den Antwortenden filtern, die seinen Wünschen weiter entsprechen, z. B. Geld überweisen. Diese Annahme lässt sich nicht auf jede mögliche Wertschöpfungskette des Angreifers übertragen. Insbesondere wenn die nachfolgenden Schritte wie die Verteilung von Schadsoftware automatisiert sind, lässt sich diese Hypothese nicht projizieren. Tatsächliche Vorgehensfehler werden von Angreifern nicht wiederholt, da sie zu einem Misserfolg des Angriffs führen oder dessen Erfolg schmälern.

Eine Cross-Site-Scripting-Verwundbarkeit in der Software *Tweetdeck* ließ sich dafür ausnutzen, Pop-up-Nachrichten bei den Nutzern zu erzeugen. Ein solches Pop-up ist in Abbildung 6 dargestellt. Da die Verwundbarkeit in einem kurzen Zeitrahmen durch die Nutzer gemeldet wurde, konnte sie zeitnah beseitigt werden. [Bra14]

3.2.4 Durch Artefakte ausgelöste transitive Effekte

Zwar sollten natürliche Elemente nicht direkt zur Messung von IT-Sicherheitsbewusstsein genutzt werden, jedoch kann eine Wechselwirkung

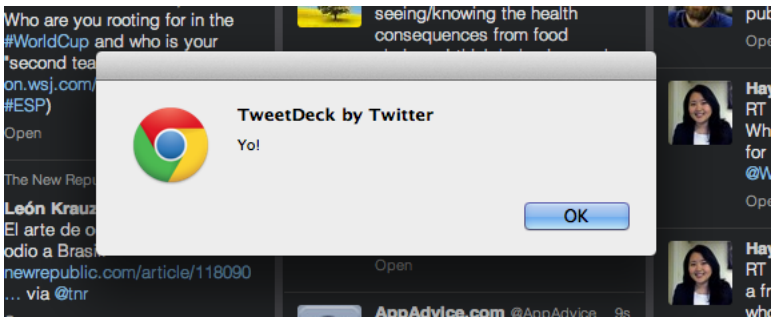


Abbildung 6: *Durch einen Cross-Site-Scripting-Angriff ausgelöstes Pop-up [Smi14].*

zwischen Artefakt und natürlichem Element einen Effekt auslösen, der eine Messung zulässt. Diese Effekte werden folgend *transitive Effekte* genannt.

Ein Beispiel für transitive Effekte sind Warnmeldungen, wie sie beispielsweise durch eingebettete Macros in den Dateiformaten von Microsoft Office erzeugt werden können [Sza14]. Die Darstellung von Artefakten und die transitiven Effekte, die durch diese ausgelöst werden, sind durch die Konfiguration der *Umgebung* bestimmt. Unterschiedliche Programme können die gleichen Daten unterschiedlich darstellen. Weiter ist die Darstellung von dem Betriebssystem und dessen Nutzerschnittstelle abhängig.

Ein weiteres aus der Forschung bekanntes Beispiel sind die Warnungen, die ein Browser generiert, wenn ein angefragter Server ein invalides Zertifikat präsentiert [Sun+09; AF13]. Es erlaubt eine Einschätzung der Vertraulichkeit und Integrität der Verbindung durch den Nutzer und damit der über sie übertragenen Daten.

3.2.5 Diskussion

Im Rahmen der vorhergehenden Abschnitte wurden mögliche Artefakte dargestellt. Diesen Artefakten ist gemein, dass sie Abweichungen des vom Nutzer *erwünschten* und *erwarteten* Verhalten des Bildschirmarbeitsplatzes als Arbeitsmittel darstellen und damit aus Nutzersicht die Integrität des Arbeitsmittels infrage stellen. Es ist die Erwartungshaltung des Nutzers, die ihn eine Ausnahmesituation feststellen lässt.

Auch muss angemerkt werden, dass eine Auflistung aller möglichen Artefakte zu keinem Zeitpunkt vollständig sein kann. Mit der technischen Fortentwicklung entwickeln sich auch die Fähigkeiten und Möglichkeiten zur Wertschöpfung des Angreifers. Diese Beobachtungen begründen die Annahme, dass jedes *unerwünschte* und potenziell *unerwartete* Verhalten als Artefakt dienen kann.

Unter dieser Annahme stellt sich jedoch die Frage nach der Güte eines Artefakts. Aus Angreifersicht ist ein Artefakt besonders gut, wenn der Nutzer die von ihm gewünschte Reaktion zeigt. Dies erreicht er beispielsweise dadurch, dass er den Nutzer täuscht, indem das Artefakt einem natürlichen Element möglichst ähnlich oder harmlos gestaltet ist, z. B. bei Phishing. Der Angreifer kann auch versuchen, eine bestimmte Handlung zu forcieren, z. B. im Falle von Ransomware. Ungeachtet dessen, wie viel *Reaktionspotenzial* ein Artefakt erzeugt, ist ein gutes Artefakt aus Sicht des Testanwenders eines, das eine valide und reliable Messung ermöglicht. Es ist an dieser Stelle nicht zu bewerten, welche Eigenschaften die Güte eines Artefaktes bezüglich des Einsatzes als Reizgeber in einer IT-Sicherheitsbewusstseinsmessung beeinflussen. Jedoch wird auch hier der Bedarf an einer Reliabilitätsbewertung der Messung hervorgehoben.

Zu den aufgegriffenen Beispielen ist anzumerken, dass einige die oben verbotene Nichtverfügbarkeit implizieren. Dies ist insbesondere immer dann der Fall, wenn der Angreifer versucht, eine Handlung des Nutzers zu forcieren, z. B. Ransomware, oder wenn eine Ressource ausgetauscht wird, z. B. Defacing. Die Eignung dieser Artefakte zur Testung ist, wie eingehend geschildert, nicht gegeben. Hier kann jedoch mit einem Kompromiss gearbeitet werden. Das Artefakt wird derart gestaltet, dass seine Auswirkungen nur temporär beobachtbar sind. Der Proband muss die Verfügbarkeit ohne signifikanten Aufwand wiederherstellen können. Damit ist dem Probanden die Möglichkeit gegeben, auf das Artefakt zu reagieren oder sein Primärziel zu bevorzugen und das Artefakt zu ignorieren. Dieses Prinzip kann beispielsweise durch eine Manipulation einer Website aufgegriffen werden, die nicht mehr beobachtbar ist, wenn der Proband die Website erneut anfragt (siehe auch „Defacing“ in Anhang A auf Seite 128).

3.3 Handlungsoptionen

Handlungsoptionen sind mögliche Reaktionen eines Nutzers auf ein Artefakt. Wie auf ein Artefakt reagiert werden kann, ist durch zwei wesentliche Faktoren bestimmt: der Gestalt des Artefakts und dem Handlungsrahmen des Probanden, der sich durch das Testumfeld ergibt. Das Artefakt bestimmt, wie mit ihm interagiert und wie es manipuliert werden kann. Eine E-Mail kann beispielsweise einen Hyperlink beinhalten, der mit einem Mausklick aktiviert werden kann (Interaktion). Die E-Mail selbst kann aber auch gelöscht werden (Manipulation).

Auch das Testumfeld kann abhängig und unabhängig von der Artefaktgestalt Handlungsoptionen anbieten. Abhängig vom Artefakt können sich beispielsweise Warnmeldungen ergeben (vgl. Abschnitt 3.2.4), mit denen der Nutzer zusätzlich konfrontiert wird, z. B. bei Makroviren (vgl. [Sza14] und Anhang A). Aber auch unabhängig von der Artefaktgestalt, kann das Testumfeld Handlungsoptionen anbieten. Werden die Bildschirmarbeitsplätze in einem Unternehmen durch dedizierte Administratoren verwaltet, wird häufig auch ein Helpdesk oder eine Hotline angeboten, an die sich ein Nutzer wenden kann, sollte dieser Unterstützung benötigen. Eine erfolgte Meldung an diese Kontaktstelle kann dann Prozesse zum Erhalt oder der Wiederherstellung der IT-Sicherheit auslösen. Diese Handlungsoption wurde bereits in der Forschung mit IT-Sicherheitsbewusstsein assoziiert [Ves11; KK06; DCF07].

Häufig geht die dedizierte Administration der Bildschirmarbeitsplätze mit stark eingeschränkten Nutzerrechten einher, was den Handlungsrahmen des Nutzers stark einschränkt. Sollten dem Nutzer hier Handlungsräume eröffnet werden, wie z. B. das bedarfsgesteuerte Auslösen des Scanvorgangs einer Host-Security-Lösung, so sind dies valide Handlungsoptionen. Je umfangreicher die dem Nutzer eingeräumten Rechte sind, desto größer ist der sich ergebende Handlungsspielraum. Im Extremfall kann die Automatisierung von Sicherheit durch einen mit entsprechenden Rechten ausgestatteten Nutzer ausgehebelt werden [WBS08, 1.1. Why automation alone is insufficient].

3.3.1 Interpretation von Handlungsoptionen in Bezug auf IT-Sicherheitsbewusstsein

Für die Konstruktion eines Maßes für IT-Sicherheitsbewusstsein gilt es zunächst, mögliche Handlungsoptionen auf ihren Bezug zur IT-Sicherheit hin zu bewerten. Handlungsoptionen lassen sich in drei Klassen einteilen – eine Handlung kann für die IT-Sicherheit *zuträglich*, *abträglich* oder *neutral* sein. Ergriffene Handlungsoptionen, die der IT-Sicherheit zuträglich sind, sind positiv zu bewerten, da ein derartiges Verhalten mit ausgeprägtem IT-Sicherheitsbewusstsein assoziiert wird. Analog repräsentiert Verhalten, das der IT-Sicherheit abträglich ist, ein Nichtvorhandensein von IT-Sicherheitsbewusstsein.

Handlungen, die keine Auswirkung auf die IT-Sicherheit haben, können nicht für eine bestimmte Ausprägung von IT-Sicherheitsbewusstsein interpretiert werden. Die zu wählende Interpretation ist abhängig von dem konkreten Artefakt. Hat der Proband eine Handlungsoption ergriffen, die durch das Artefakt angeboten wird, ist davon auszugehen, dass der Proband das Artefakt wahrgenommen hat. Ist das Artefakt als solches erkennbar und gleicht keinem natürlichen Element, kann diese Wahrnehmung nicht der Erwartungshaltung des Probanden entsprechen. Unter diesen Umständen repräsentiert das Ergreifen einer Handlungsoption, die dem Artefakt entspringt, ein Fehlen von IT-Sicherheitsbewusstsein, auch wenn die Handlungsoption für die IT-Sicherheit von neutraler Bedeutung ist.

Es ist denkbar, dass ein Nutzer IT-sicherheitsbewusst ist, aber eine Handlungsoption ergreift, die der IT-Sicherheit abträglich ist. Stanton u. a. ordnen IT-sicherheitsrelevantes Verhalten in zwei Dimensionen an: Expertise und Intention [Sta+05]. So ist es möglich, dass ein IT-sicherheitsbewusster Nutzer absichtlich der IT-Sicherheit abträglich handelt. IT-Sicherheitsbewusstsein bezieht, wie in Abschnitt 2.1.1 geschildert, auch motivatorische Aspekte mit ein. Es existiert eine Fehlstellung, die so zum Ausdruck kommt. Ebenso ist denkbar, dass der Nutzer nicht über die nötige Expertise verfügt und ohne erkennbaren Anlass eine der IT-Sicherheit zuträgliche Handlungsoption ergreift. Dies ist im Rahmen einer Messungenauigkeit hinzunehmen. Um diese Messungenauigkeit zu reduzieren, kann die artefaktbasierte IT-Sicherheitsbewusstseinsmessung mit einer wissensbasierten Befragung kombiniert werden.

Neben der Einteilung in diese drei Klassen ergibt sich Potenzial für eine weitere Gewichtung der Handlungsoptionen. So können durch eine Handlungsfolgenabschätzung die Auswirkungen der Handlungsoptionen auf die IT-Sicherheit abgeschätzt, entsprechend bewertet und geordnet werden. Diese Abschätzung ist den Gegebenheiten des Umfelds, in dem getestet wird, gegenüber nicht agnostisch. Sie müsste für jedes Umfeld erneut erstellt oder angepasst werden. Die Handlungsfolgenabschätzung ist mit signifikantem Aufwand verbunden. Der rudimentäre klassenbasierte Ansatz ist weitestgehend unabhängig, kann auf ein beliebiges Umfeld übertragen werden und ist aus diesem Grund zu favorisieren.

3.4 Masskonstruktion

Im Rahmen dieses Unterkapitels wird ein Maß hergeleitet, das erlaubt, die erfassten Nutzerreaktionen zu einem Maß für IT-Sicherheitsbewusstsein zusammenzufassen. Dazu wird ein Maß für das *individuelle* IT-Sicherheitsbewusstsein formuliert. Anschließend werden Optionen zur *Skalierung* dieses Maßes auf Gruppen vorgestellt und diskutiert.

Ein Maß sollte vor allem nützlich sein. Dafür muss es aussagekräftig sein. Das wird erreicht, indem sich der erhaltene Wert in einen Referenzrahmen einbetten lässt. Dazu muss der Wertebereich des Maßes ein bekanntes Minimum und Maximum besitzen. Darüber hinaus ist es wünschenswert, dass die Berechnungsvorschrift semantisch ähnlich zu bekannten und etablierten Maßen ist.

3.4.1 IT-Sicherheitsbewusstsein von Individuen

In Abschnitt 2.2 wurde dargestellt, dass bei Experimenten verschiedene Interaktionsformen mit Artefakten erfasst wurden. Abschnitt 3.3 ist zu entnehmen, dass sich die relevanten Handlungsoptionen in zwei Kategorien einteilen lassen: der IT-Sicherheit zuträglich und der IT-Sicherheit abträglich. Obwohl Gréaux mit dem Aufgreifen der chronologischen Ereignisabfolge in dem Beobachtungszeitraum [Gré15] einen interessanten alternativen Ansatz aufzeigt, wird zur besseren Vergleichbarkeit mit existierender Forschung (siehe dazu Abschnitt 2.2) ein streng ereignisorientierter Ansatz verfolgt.

Werden die Bildschirmarbeitsplätze zentral administriert, sind die Handlungsoptionen, die der Nutzer wahrnehmen kann, limitiert (vgl. Abschnitt 3.3). Dies ist insbesondere wahr für die IT-Infrastruktur des ITS.APT-Anwendungspartners (vgl. Abschnitt 1.2). Im allgemeinen Fall können mehrere Handlungsoptionen, je nach eingesetztem Artefakt, in die Klasse der IT-Sicherheit zuträglichen Handlungsoptionen fallen. Für die weitere Betrachtung wird vereinfachend angenommen, dass die Meldung die einzige der IT-Sicherheit zuträgliche Handlungsoption ist. Sie wird mit *r* für „*report*" notiert. In Abschnitt 3.3 wird ebenfalls dargestellt, dass jede Artefaktinteraktion der IT-Sicherheit abträglich ist, und wird mit *i* für „*interaction*" notiert. Die erfassten Handlungen als Ergebnis eines Tests präsentieren sich als Quadrupel (vgl. Abbildung 7):

```
[proband, artefakt, zeitpunkt, handlungsoption]
```

Abbildung 7: *Datenmodell der erfassten Reaktionen auf Artefakte.*

Pro Proband und Test sollte nur ein Artefakt zum Einsatz kommen. Es ist zwar die theoretische Option gegeben, dass sich ein Proband gleichzeitig in mehreren Tests befindet, doch ist diese Situation nicht erstrebenswert. Ein Artefakt wird zusätzlich in die Situation eingegeben. Die Umgebungsvariablen wie der Stress am Arbeitsplatz oder seine Variabilität sollten während der Testung der tatsächlichen, gewohnten Arbeitssituation möglichst ähnlich sein. Jeder Test wirkt durch das Artefakt als Reizgeber auf den Probanden ein. Befindet sich der Proband gleichzeitig in zwei Tests, beeinflussen sich diese gegenseitig. Dies ist unerwünscht und es gilt diese Situation zu vermeiden.

Ein Proband kann jedoch mehrere Handlungsoptionen ergreifen. So ist es denkbar, dass ein Nutzer in einem Test sowohl mit dem Artefakt interagiert als auch danach Kontakt mit dem Helpdesk aufnimmt, um das Artefakt zu melden.

Die chronologische Ordnung der Ereignisse lässt sich hier bei der Maßkonstruktion vernachlässigen. Funktionieren die IT-Sicherheitsprozesse der administrierten IT-Infrastruktur, macht es keinen Unterschied für die Sicherheit der IT-Infrastruktur, ob der Nutzer zuerst mit dem Artefakt interagiert hat und anschließend das Helpdesk auf das Artefakt aufmerksam macht, oder der Nutzer zuerst die Konfrontation mit dem Artefakt meldet und danach doch noch mit diesem interagiert. In jedem Fall ist das Artefakt durch das Helpdesk aus der Perspektive der

IT-Sicherheit zu bewerten und entsprechend zu verfahren. Der Nutzer hat sein IT-Sicherheitsbewusstsein durch die Meldung gezeigt. Aus diesen Betrachtungen ergeben sich vier unterschiedliche Kombinationen der Handlungsoptionen, die Ereignisklassen. Tabelle 2 zeigt eine Übersicht.

Tabelle 2: *Kombinationen aus Handlungsoptionen der Tests nach ihren Ereignisklassen. $\{\varnothing\}$ notiert das leere Ereignis, $\{i\}$ eine Interaktion mit dem Artefakt und $\{r\}$ einen Report des Artefakts. Das Ausbleiben einer Reaktion dieses Typs ist mit der Negation $\neg$ gekennzeichnet.*

Notation	Ereignisklasse
$\{i \wedge \neg r\}$	Mit dem Artefakt interagieren und das Helpdesk nicht informieren.
$\{i \wedge r\}$	Mit dem Artefakt interagieren und das Helpdesk informieren.
$\{\neg i \wedge r\}$	Nicht mit dem Artefakt interagieren und das Helpdesk informieren.
$\{\neg i \wedge \neg r\} \mid \{\varnothing\}$	Nicht mit dem Artefakt interagieren und das Helpdesk nicht informieren. In diesem Fall zeigt der Proband keine Reaktion.

Interagiert der Proband nicht mit dem Artefakt, meldet dieses jedoch dem Helpdesk, demonstriert der Proband sein IT-Sicherheitsbewusstsein. Auch in dem Fall, in dem der Proband mit dem Artefakt interagiert, es aber dem Helpdesk meldet, demonstriert der Proband sein IT-Sicherheitsbewusstsein. Er hat das Artefakt als solches erkannt und die erwünschte Handlungsoption ergriffen. Interagiert ein Proband mit dem Artefakt und meldet dieses nicht dem Helpdesk, ist dies als Beleg für sein fehlendes IT-Sicherheitsbewusstsein zu interpretieren. Der Proband hat das Artefakt wahrscheinlich wahrgenommen, sonst könnte er nicht mit diesem interagieren. Er hat dieses aber entweder nicht als solches erkannt oder sich trotz der Tatsache, dass er es erkannt hat, gegen eine Meldung entschieden. Wenn der Proband weder mit dem Artefakt interagiert noch es dem Helpdesk meldet, ist ohne aufwendiges Monitoring der Arbeitsumgebung nicht zu ermitteln, ob der Proband das Artefakt überhaupt wahrgenommen hat. In diesem Fall wird zugunsten der Probanden aufgelöst. Es wird angenommen, dass der Proband das Artefakt wahrgenommen und erkannt hat und der Proband

aus diesem Grund nicht mit dem Artefakt interagiert hat. Das Konzept ist in Abbildung 8 illustriert.

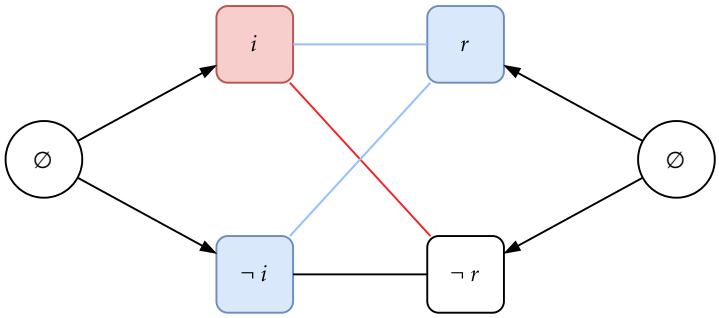


Abbildung 8: *Kombinationen möglicher Reaktionstypen von Probanden auf Artefakte. {∅} notiert das leere Ereignis, {i} eine Interaktion mit dem Artefakt und {r} einen Report des Artefakts. Das Ausbleiben einer Reaktion dieses Typs ist mit der Negation ¬ gekennzeichnet. Kanten fassen die Ereignisse zu Klassen zusammen. Wertungen sind farblich hervorgehoben. Eine Klasse, die nur rote, aber keine blauen Ereignisse enthält, wird als der IT-Sicherheit abträglich bewertet. Enthält eine Ereignisklasse ein blaues Ereignis, wird die Klasse als der IT-Sicherheit zuträglich interpretiert.*

Mit dieser Interpretation ergibt sich als Maß für das individuelle IT-Sicherheitsbewusstsein eines Probanden s die geschätzte Wahrscheinlichkeit P des Gegenereignisses des unerwünschten Ereignisses, $\{i \wedge \neg r\}$. Der wahre Wert der Wahrscheinlichkeit für $P\{i \wedge \neg r\}$ ist unbekannt, lässt sich aber durch die Anwendung des Tests für mehrere Artefakte eines Sets A approximieren. Gleichung (3.1) zeigt die sich ergebende Verrechnungsvorschrift.

$$Individual\ Awareness(s) = 1 - P_{A,s}(\{i \wedge \neg r\}) \text{ für einen Probanden } s \qquad (3.1)$$

Das so gestaltete Maß hat ein Minimum von 0 und ein Maximum von 1. Bei Messungen, in deren Umfeld der Nutzer keine Option hat, eine der IT-Sicherheit zuträgliche Handlungsoption zu ergreifen, ergibt sich die Schätzung dieser Wahrscheinlichkeit aus dem invertierten Anteil der Probanden, die mit dem Artefakt interagiert haben. Damit stellt der Wert eine generische Form der *click rate* dar. Wird die Interaktion mit dem Artefakt nicht erfasst, ergibt sich der Wert über den Anteil der Probanden, die das Artefakt gemeldet haben, die *report rate*.

Dies ist konsistent zu den in der verwandten Forschung genutzten Maßzahlen (vgl. Abschnitt 2.2.4 auf Seite 22).

3.4.2 IT-Sicherheitsbewusstsein von Gruppen

Zur Skalierung auf eine Gruppe von Probanden wie der Belegschaft eines Unternehmens bieten sich zwei mögliche Optionen an. Zum einen kann der Wert über alle Probanden gemittelt werden. Diese Vorgehensweise ist insbesondere zur Validitätsprüfung nötig, da hier der Bezug zu den verwandten Arbeiten genutzt wird. Die zweite Option ist modellgetrieben.

Für den Erhalt der IT-Sicherheit einer IT-Infrastruktur ist lediglich notwendig, dass ein Nutzer das IT-sicherheitsrelevante Ereignis meldet oder kein Nutzer mit dem Artefakt interagiert. Wird dabei das in Gleichung (3.1) dargestellte Prinzip befolgt, ergibt sich die Wahrscheinlichkeit für das Gegenereignis (ein Proband interagiert, kein Proband meldet das Ereignis) einer Gruppe aus n Probanden wie folgt:

$$P(\{i \wedge \neg r\}) \cdot P(\{\neg r\})^{n-1} \tag{3.2}$$

Die einfache Invertierung des Wertes aus Gleichung (3.2) wie in Gleichung (3.1) ist jedoch nicht zielführend. Es impliziert die Annahme, dass die gesamte Belegschaft die Gelegenheit hat, das Ereignis zu melden. Wie viele Individuen der Gruppe einen potenziellen Angriff beobachten können, ist unbekannt und wahrscheinlich variabel. Um den Umfang der Beobachtbarkeit zu modellieren, wird die Funktion $\omega\colon x \to [0,1]$ genutzt. Damit kann eine Verteilung genutzt werden, um die Wahrscheinlichkeit zu modellieren, dass x Individuen ein Artefakt beobachten können. Soll keine Annahme über diese Wahrscheinlichkeit getroffen werden, muss die Gleichverteilung für alle möglichen Angriffsgrößen angenommen werden. In diesem Fall kann $\omega(x) = 1/n$ gesetzt werden. Der Term aus Gleichung (3.2) wird mit $\omega(x)$ multipliziert. Wird der Term anschließend für alle möglichen $x = 1,\ldots,n$ aufsummiert und von 1 abgezogen, ergibt sich die in Gleichung (3.3) dargestellte Rechenvorschrift. Auch hier muss die Wahrscheinlichkeit mit einem Set von Artefakten A approximiert werden.

$$Group\ Awareness(n) = 1 - \frac{1}{n} \cdot \sum_{x=1}^{n} P_A(\{i \wedge \neg r\}) \cdot P_A(\{\neg r\})^{x-1} \tag{3.3}$$

Für ein $n = 1$ ergibt sich aus der in Gleichung (3.3) dargestellten Berechnungsvorschrift das individuelle IT-Sicherheitsbewusstsein (vgl. Gleichung (3.1)). Diese modellgetriebene Skalierung ist zur Bewertung der Gruppe der Nutzer nützlicher als die Skalierung über den Mittelwert des individuellen IT-Sicherheitsbewusstseins, da sie eine realitätsnähere Aussage über den Sicherheitsgewinn einer Nutzerschaft erlaubt.

3.5 Fazit

In diesem Kapitel wurde das *Situationsmodell* vorgestellt. Darauf aufbauend wurden die Konzepte des *Artefakts* und der *Handlungsoption* vorgestellt. Artefakte sind künstlich in die Situation eingebrachte Elemente mit IT-Sicherheitsbezug und fungieren im Rahmen eines Tests als Reizgeber. Die Einbettung der Situation in das Testumfeld bestimmt die konkrete Ausprägung der Artefakte und die dem Probanden zur Verfügung stehenden Handlungsoptionen. Auf Basis dieser Konzepte wurde ein Maß entwickelt, das eine Interpretation ergriffener Handlungsoptionen, und damit eine Bewertung des IT-sicherheitsrelevanten Verhaltens eines Individuums oder einer Gruppe aus Individuen, zulassen.

Die Idee einer Leistungsbewertung des IT-sicherheitsrelevanten Verhaltens wurde bereits von Vroom & von Solms diskutiert, aber als zu aufwendig betrachtet [VvS04]. Sie beschreiben die möglichen Auswirkungen von individuellen und organisatorischen Einflüssen auf die Validität und Reliabilität der Bewertung. Um diesen Effekten entgegenzuwirken und einen verlässlichen Messwert zu erzielen, muss die Messung mehrfach durchgeführt werden. Dies wird durch folgende Bedingung 1 widergespiegelt.

Bedingung 1: *Ein Proband muss mehrere Tests mit unterschiedlichen Artefakten durchlaufen.*

Damit die Messung ökonomisch ist, muss die Messung voll automatisiert werden. Dazu wird ein Werkzeug entwickelt, das die erforderlichen Funktionalitäten implementiert. Grundlegend ergibt sich zunächst Anforderung 1:

Anforderung 1: *Das Werkzeug muss Artefakte ausbringen können.*

Diese Artefakte sind zum Teil zu personalisieren, um den Probanden gezielt anzusprechen (vgl. Abschnitt 3.2.1). Damit ergibt sich Anforderung 2:

Anforderung 2: *Das Werkzeug muss Artefakte personalisieren können.*

Um den Test vollständig abzubilden, müssen die ergriffenen Handlungsoptionen aufgezeichnet werden und diese auch interpretiert werden. Dies spiegelt sich in Anforderung 3 wider:

Anforderung 3: *Das Werkzeug muss in der Lage sein, alle relevanten Handlungsoptionen zu monitoren.*

Die durch die Probanden ergriffenen und durch das Werkzeug aufgezeichnete Handlungsoptionen ist zur Skalierung durch die oben vorgestellten Maße zu bewerten. Es ergibt sich Anforderung 4:

Anforderung 4: *Das Werkzeug muss die durch die Probanden ergriffenen Handlungsoptionen auswerten.*

Die im Rahmen dieses Kapitels vorgeschlagene Generalisierung der Phishing-Feldversuche, wie in Abschnitt 2.2.4 (Seite 22) beschrieben, adressiert die Schwächen eben dieser Messmethode (vgl. Abschnitt 2.3). Die größere Bandbreite an erfassbarem Verhalten ermöglicht eine höhere Validität der Aussage in Bezug auf das IT-Sicherheitsbewusstsein. Durch das vorgestellte Maß ist eine Skalierung gegeben. Eine Reliabilitätsbewertung ist an dieser Stelle ausstehend. Diese kann jedoch über die höhere mögliche Variabilität in den zur Messung nutzbaren Artefakten einfacher durchgeführt werden. Sie muss Teil der späteren Erprobung sein. Die artefaktbasierte Messung von IT-Sicherheitsbewusstsein lässt sich wie die Phishing-Feldversuche automatisiert durchführen.

4 Analyse der ethischen und juristischen Rahmenbedingungen

Im Rahmen eines Feldexperiments sind ethische und juristische Fragestellungen implizit. Damit die artefaktbasierte Messung von IT-Sicherheitsbewusstsein Adaption über diese Arbeit hinaus erfahren kann, ist es unabdingbar, die Anforderungen, die sich aus den ethischen und juristischen Fragestellungen ergeben, strukturiert zu erfassen. Im weiteren Entwurf des Werkzeuges und der geplanten Studie, die in den folgenden Kapiteln thematisieren werden, können die herausgearbeiteten Rahmenbedingungen so adäquat adressiert werden.

Zunächst werden Arbeiten vorgestellt und diskutiert, die ethische Fragestellungen zu Feldversuchen im Bereich der Forschung zu Phishing thematisieren. Diesen lässt sich jedoch nur eingeschränkt Anleitung zum Einsatz der artefaktbasierten Messung von IT-Sicherheitsbewusstsein entnehmen. Eine Selbstbeurteilung des Vorhabens erlaubt die strukturierte Erfassung der ethischen und juristischen Rahmenbedingen. Diese ist in Abschnitt 4.2 dokumentiert. Anforderungen werden entsprechend abgeleitet. Das Kapitel schließt mit einem Fazit.

4.1 Verwandte Arbeiten

Es existieren bereits Beiträge innerhalb des akademischen Diskurses zur Umsetzung ethisch unbedenklicher Feldversuche im Bereich der Phishing-Experimente. Da die im Rahmen dieser Arbeit entwickelte Methode Phishing-Experimente verallgemeinert, werden diese Beträge in diesem Abschnitt diskutiert.

Jakobsson & Ratkiewicz erstellten die erste Arbeit, die sich spezifisch mit der Frage der Ethik in Phishing-Feldexperimenten auseinandersetzt [JR06]. In ihrem Versuchsaufbau fälschen sie Nachrichten innerhalb des Nachrichtensystems

einer Online-Auktionsplattform. Die Autoren postulieren ein Experiment als „ethisch“, wenn dessen Probanden nicht zu Schaden kommen. Sie stellen das Prinzip des ethischen Experiments einem „akkuraten“ Experiment gegenüber. Ein Experiment ist akkurat, wenn seine Erfolgswahrscheinlichkeit mit der Erfolgswahrscheinlichkeit eines echten Angriffs vergleichbar ist. Ob die Teilnehmer dieses Experiments über ihre Teilnahme aufgeklärt wurden, ist nicht dokumentiert. Möglicherweise wurden sie getäuscht. Auch eine Befugnis durch die Betreiber der Online-Auktionsplattform ist nicht überliefert. Auch wenn den Probanden kein monetärer Schaden durch die Durchführung der Studie entstanden ist, widerspricht dieses Vorgehen anerkannten Richtlinien zur Durchführung von Humanexperimenten (vgl. Abschnitt 4.2).

Finn & Jakobsson geben ein Jahr später Einblick in den Entscheidungsprozess einer Ethikkommission in Hinblick auf Phishing-Experimente [FJ07]. Insbesondere werden Studien nach den Prinzipien des Belmont-Report [Res78] bewertet. Es wird argumentiert, dass allein das Wissen der Probanden um die Studie das Verhältnis zwischen dem durch den Probanden wahrgenommenen Risiko und dem tatsächlichen Risiko verschiebt. Die Autoren sprechen sich damit gegen die informierte Einwilligung der Probanden aus. Darüber hinaus stellen die Autoren eine Nachbesprechung mit den Probanden infrage. Die zugrunde liegenden Beobachtungen basieren auf einem bestimmten Phishing-Experiment [Jag+07; JF08]. In diesem reagierten die Probaden missgünstig auf die Ansprache zur Nachbesprechung. Da sowohl Form als auch Wortlaut der Probandenansprache nicht überliefert sind, könnte hier von einem Einzelfall ausgegangen werden [BPJ13].

Buchanan u. a. [Buc+11] diskutieren Implikationen der Forschung auf dem Bereich der IT-Sicherheit für die Bewertung durch eine Ethikkommission. Sie führen den Belmont-Report [Res78] als etablierten Leitfaden für die Bewertung der Ethik klassischer Humanexperimente an. Keines der drei diskutierten Beispiele ist mit dem im Rahmen dieser Arbeit geplanten Vorgehen vergleichbar. Vielmehr wird der Begriff des Probanden als Erweiterung der zu erforschenden Technik genutzt. Die Autoren unterstreichen damit das Verständnis von IT-Infrastruktur als soziotechnologisches Konstrukt.

Busch u. a. schlugen 2016 ein Experiment vor, um gezielt mögliche negative Effekte der Täuschung auf die Probanden eines Phishing-Experiments zu

erfassen [Bus+16]. Eine Dokumentation der Durchführung des Experiments ist nicht publik.

Die einschlägige Literatur zu ethischen Feldversuchen liefert nur sehr eingeschränkt Anleitung, die sich auf die hier geplante Studie übertragen lässt. Sie verweist jedoch auf den Belmont-Report [Res78] als Grundlage ethischer Bewertungen von Humanexperimenten. Dieser bildet die Grundlage für die Selbstbeurteilung.

4.2 Selbstbeurteilung

Eine ethische Bewertung des Einsatzes der artefaktbasierten IT-Sicherheitsbewusstseinsmessung im Klinikumfeld (vgl. Abschnitt 1.2) erlaubt Rückschlüsse auf die zu beachtenden Rahmenbedingungen. Diese sind sowohl bei der Gestaltung des Werkzeuges als auch bei der Durchführung des Experiments zu beachten. Anhand etablierter Leitfäden wird im Rahmen dieses Unterkapitels zu diesem Zweck eine Selbstbeurteilung durchgeführt.

Der Menlo-Report [KD12] schlägt einen ethischen Leitfaden für die Computer- und Informationssicherheitsforschung vor. Er bietet zu diesem Zweck vier grundlegende Prinzipien an, die Entscheidungen bei der Implementierung von Experimenten leiten sollen. Damit spezifiziert der Menlo-Report die drei Prinzipien des anerkannten Belmont-Report [Res78] und erweitert diese mit einem vierten Prinzip.

Der Menlo-Report lässt damit eine eigene Aufbereitung und die umfängliche Ableitung von Anforderungen an das zu entwickelnde Werkzeug und seine Erprobung im KRITIS-Umfeld zu. Die vier Prinzipien des Menlo-Reports lauten:

1. **„Respect for Persons“:** Die Teilnahme an dem Forschungsprojekt ist freiwillig und erfolgt nach informierter Einwilligung. Individuen sind als autonome Vertreter ihres eigenen besten Interesses zu sehen. Personen mit eingeschränkter Autonomie (z. B. Kinder oder andere vermindert geschäftsfähige Personen) steht gesonderter Schutz zu.
2. **„Beneficence“:** Füge keinen Schaden zu. Maximiere wahrscheinlichen Gewinn und minimiere wahrscheinlichen Schaden.

3. **„Justice“:** Jede Person verdient Gleichbehandlung. Die Probandenselektion soll fair und die Last unter allen Probanden verteilt sein.
4. **„Respect for Law and Public Interest“:** (Rechtliche) Sorgfaltspflicht ist geboten. Ebenso Transparenz bezüglich der Methodik und der Ergebnisse. Verantwortungspflicht ist gefordert.

Um die Würdigung dieser Prinzipien zu ermöglichen, ist es unabdingbar, alle beteiligten Parteien und ihre Bedürfnisse zu erfassen. Diese werden im Anschluss im Hinblick auf die vier oben genannten Prinzipien ausgewertet.

4.2.1 Stakeholder des Experiments

Im Rahmen des Experiments werden im laufenden klinischen Betrieb Bildschirmarbeitsplatznutzern Artefakte präsentiert und deren Reaktion auf diese Artefakte aufgezeichnet. Bei einem derartigen Experiment existieren viele Interessensträger (engl. *Stakeholder*). Insbesondere sind in der folgenden Betrachtung jene Interessensträger fokussiert, aus deren Interessen sich Anforderungen oder Bedingungen ableiten lassen. Damit ergeben sich die folgenden Personengruppen als Stakeholder:

Probanden Nutzer eines Bildschirmarbeitsplatzes der IT-Infrastruktur, denen Artefakte präsentiert und deren Reaktionen aufgezeichnet werden, sind *Probanden*. Als Angestellte des *Arbeitgebers* sind sie diesem gegenüber verpflichtet und ihre Interessen werden vor diesem durch die *Mitarbeitervertretung* repräsentiert. In dieser Rolle wird diese Personengruppe durch den Gesetzgeber geschützt, dies wird in Abschnitt 4.2.5 aufgegriffen. Insbesondere ist einer möglichen negativen Auswirkung durch den Arbeitgeber auszuschließen. Damit einher geht auch der Bedarf nach der vertraulichen Behandlung der Identität jedes Probanden. Im Rahmen dieses Experiments sollen die Probanden Reaktionen auf Artefakte während des regulären betrieblichen Ablaufs zeigen. Diese Artefakte können den Handlungsablauf der Probanden unterbrechen und damit die Zielerreichung beeinträchtigen. Es ist davon auszugehen, dass diese Unterbrechung Stress erzeugt. Probanden sind daran interessiert, keinem unnötigen Stress ausgesetzt zu sein. Ein Proband möchte selbstbestimmt handeln. Das betrifft sowohl die Teilnahme an dem Experiment als auch die Verwertung der erfassten Daten.

Nichtprobanden *Nichtprobanden* sind Nutzer und Nutznießer der IT-Infrastruktur, in der getestet wird, die aber nicht an dem Experiment partizipieren. Sie sind durch eine mögliche Störung des Betriebs der IT-Infrastruktur direkt betroffen. Dies trifft insbesondere auf Patienten zu. Diese sind von eventuellen Störungen im Betrieb direkt betroffen und Auswirkungen sind unter Umständen katastrophal. Darüber hinaus haben Nichtprobanden einer Durchführung des Experiments nicht zugestimmt. Die Gruppe der Nichtprobanden ist unkontrolliert. Insbesondere können sich unter ihnen Kinder oder andere vermindert geschäftsfähige Personen befinden, denen ein besonderer Schutzbedarf zuzumessen ist. Aus diesen oben genannten Gründen ist ein besonderer Schutzbedarf für Nichtprobanden festzustellen. Durch die Höhe des möglichen Schadens muss, ungeachtet der geringen Eintrittswahrscheinlichkeit, ein besonderes Risiko anerkannt werden.

Betreiber der Infrastruktur Betreiber und Eigner der IT-Infrastruktur, die IT-Abteilung und das Helpdesk, sind besonders an dem störungsfreien Betrieb der IT-Infrastruktur interessiert. Die Betreiber der Infrastruktur sind dem Wohl aller Nutzer und Nutznießer der IT-Infrastruktur, Proband oder nicht, verpflichtet. Darüber hinaus ist der Ressourcenaufwand, der durch eine Testung erzeugt wird, zu minimieren.

Arbeitgeber Der Arbeitgeber ist an einer besonders *guten* Bewertung für seine Mitarbeiter interessiert. Er ist in der Position Mitarbeiter mit einem *schlechten* Ergebnis zu sanktionieren. Es ist daher sicherzustellen, dass die Ergebnisse des Experiments ausschließlich anonymisiert kommuniziert werden. Sollten die Ergebnisse des Experiments missverständlich, insbesondere negativ konnotiert, kommuniziert werden, fürchtet der Arbeitgeber einen Reputationsschaden.

Mitarbeitervertretung Die *Mitarbeitervertretung*, im Falle der Klinik in Form des Personalrates, sichert die Rechte der Mitarbeiter und vertritt deren Interessen. Der Personalrat wird durch die Mitarbeiter des Unternehmens aus den Mitarbeitern selbst gewählt. Sie vertritt damit Probanden und Nichtprobanden gleichermaßen.

4.2.2 „Respect for Persons“

Die Anwendung des Prinzips „Respekt für Personen“ ist mehrschichtig. Es adressiert alle involvierten Personen. Zunächst wird das Prinzip für die Gruppe der Probanden und anschließend für alle weiteren Stakeholder angewendet.

Probanden

Potenziellen Probanden muss ein höchstmögliches Maß an Selbstbestimmung gewährt werden. Dies wird typischerweise über eine Vorgehensweise ermöglicht, die das *informierte Einverständnis* der potenziellen Probanden vor der Teilnahme an dem Experiment einholt.

Bedingung 2: *Kann ein informiertes Einverständnis, beispielsweise durch eine verringerte Geschäftsfähigkeit (vgl. [BGB20, § 104–107]), nicht erteilt werden, so sind diese potenziellen Probanden von der Teilnahme auszuschließen.*

Bedingung 3: *Die Erteilung des Einverständnisses muss ohne negative Auswirkungen für den potenziellen Probanden ablehnbar sein.*

Wie bereits in Kapitel 2 und Abschnitt 4.1 dargestellt, ist das Nutzerverhalten der Probanden sensitiv gegenüber einer Aufklärung. Aus diesem Grund können Probanden nicht vor der Durchführung über das Experiment informiert werden. Der frühestmögliche Zeitpunkt zur Information der Probanden ist nach der Datenerhebung. Zu diesem Zeitpunkt müssen Probanden umfänglich informiert werden. Darüber hinaus muss den Probanden die Möglichkeit zum Widerruf der Teilnahme angeboten werden.

Bedingung 4: *Die Probanden müssen frühestmöglich über das Experiment, ihre Teilnahme, ihre Rechte und wie sie diese in Anspruch nehmen können informiert werden.*

Da es sich bei den Probanden um Angestellte eines Unternehmens handelt, ergibt sich der Umstand, dass die Probanden in dieser Rolle durch eine *Mitarbeitervertretung*, den Betriebs- oder Personalrat, vertreten werden. Auch wenn die Probanden nicht vor der Datenerhebung aufgeklärt werden können, kann der gewählte Vertreter der Probanden deren Rechte wahren. Dazu muss der Betriebs- oder Personalrat vollumfänglich aufgeklärt werden und seine

Zustimmung zu dem geplanten Vorhaben erteilen. Diese Zustimmung ist durch eine geschlossene Betriebs- oder Dienstvereinbarung zu dokumentieren.

Der Betriebs- oder Personalrat ist verpflichtet, die durch ihn geschlossene Betriebs- oder Dienstvereinbarung im Unternehmen zu publizieren [Bet20, § 77 Abs. 2]. Das Gesetz fordert allerdings nicht die persönliche Ansprache der betroffenen Mitarbeiter. Im Falle des Anwendungspartners, dem UKSH, werden die getroffenen Dienstvereinbarungen im Intranet abgelegt. Um die Möglichkeit einer Verzerrung der Messung durch Anforderungsmerkmale so gering wie möglich zu halten, ist die Betriebs- oder Dienstvereinbarung so generisch wie möglich zu halten. Hier ergibt sich ein Kompromiss. Zwar ist hier nicht von einer informierten Einwilligung der Probanden zu sprechen, jedoch von der informierten Einwilligung der gewählten Vertreter.

Nichtprobanden

Neben den Probanden existiert die Gruppe der Nichtprobanden. Sie sind von dem Betrieb der IT-Infrastruktur abhängig und lassen sich in zwei Untergruppen aufteilen.

Die erste Untergruppe sind Personen, die nicht an dem Experiment beteiligt sind. Dies sind insbesondere Angestellte, die nicht als Proband ausgewählt wurden, sowie Nutznießer der kritischen Infrastruktur. Diese Gruppe darf durch die Durchführung des Experiments nicht beeinträchtigt werden.

Bedingung 5: *Durch den Einsatz des Werkzeugs im Rahmen des Experiments darf kein Nichtproband beeinträchtigt werden, der nicht vorher explizit dem Vorhaben zugestimmt hat.*

Bedingung 5 wird insbesondere durch die Probandenselektion (vgl. Abschnitt 6.1.1) und ein Nutzertrackingkonzept (vgl. Abschnitt 5.3) adressiert. Nur durch Letzteres kann gewährleistet werden, dass Artefakte ausschließlich den Probanden präsentiert werden.

Ein weiteres Argument gegen das Einholen eines informierten Einverständnisses der Probanden in diesem Feldversuch basiert auf der Kombination von Bedingung 5 und den Beobachtungen von Finn & Jakobsson [FJ07]. Es kann nicht nur zu einer Verzerrung des wahrgenommenen Risikos der Probanden

bezüglich des Messinstrumentes und damit des Artefakts kommen. Wenn der Proband einen tatsächlichen Angriff während des Testzeitraums wahrnimmt, kann dieser irrtümlicherweise als Teil der Testung angenommen werden. Dies kann die Sicherheit der gesamten IT-Infrastruktur beeinflussen.

Die zweite Untergruppe sind die Mitarbeiter der Organe des IT-Betriebs, die mit den Probanden während des Experiments in Kontakt kommen. Dies trifft insbesondere auf das Helpdesk zu. Es ist zu erwarten, dass Probanden während der Testung die gesamte gebotene Breite möglicher Handlungsoptionen ausnutzen werden. In diesem Zusammenhang kann es zu einem erhöhten Aufkommen von Nutzeranfragen am Helpdesk kommen. Auch dürfen Nutzeranfragen, die durch einen Test ausgelöst werden, nicht mit tatsächlichen möglichen Sicherheitsvorfällen in dem gleichen Zeitraum verwechselt werden. Das Helpdesk ist wie alle Teile des IT-Betriebs vorab über die Testung zu informieren und die Durchführung ist entsprechend abzustimmen.

4.2.3 „Beneficence“

Zur Demonstration der Wohltätigkeit ist der durch die Durchführung des Experiments erzielte Mehrwert den Risiken der Durchführung und der Zumutbarkeit gegenüber der Probanden (vgl. Abschnitt 2.2.1) abzuwägen. Die Ziele des Experiments sind die Demonstration der Durchführbarkeit derartiger Experimente sowie der Funktionsfähigkeit des entwickelten Werkzeugs in Bezug auf Validität, Reliabilität und Gebrauchstauglichkeit. Damit wird der Weg für die systematische Erfassung der Auswirkungen von IT-sicherheitsbewusstseinserhöhenden Interventionen auch in den empfindlichsten Teilen unserer Gesellschaft, den kritischen Infrastrukturen, bereitet (vgl. Kapitel 1).

Dem entgegen stehen verschiedene Risiken. Insbesondere sind die hohen Risiken für Nichtprobanden, die sich bis zur Gefahr für die körperliche Unversehrtheit erstrecken können. Diese Risiken sind unbedingt zu minimieren, was bereits in Bedingung 5 festgehalten ist. Darüber hinaus ist das Risiko einer Störung der IT-Infrastruktur durch einen Fehler des Werkzeugs zu minimieren. Bedingung 6 wird durch die ausgiebige Testung des Werkzeugs selbst (vgl. Abschnitt 5.2.1) adressiert:

Bedingung 6: *Die Wahrscheinlichkeit für eine Beeinträchtigung der IT-Infrastruktur durch den Einsatz des Werkzeugs ist zu minimieren.*

Für die Probanden ergibt sich durch die Präsentation der Artefakte eine zusätzliche Belastung. Diese ist integraler Bestandteil des Experiments und als solcher unvermeidbar. Situationen, in denen Mitarbeiter eines Unternehmens einen Angriff auf die IT-Infrastruktur wahrnehmen, sind jedoch erwartbar (vgl. Abschnitt 3.2) und können durch ihre Art nicht als unverhältnismäßig angesehen werden. Während der Datenerhebung kann durch die Häufung der Artefaktpräsentationen eine erhöhte Belastung für die Probanden entstehen. Aus diesem Grund muss die Belastung idealerweise so verteilt werden, dass die Probanden durch die Testung nicht beeinträchtigt werden. Dem wird bereits durch die konzeptionelle Beschränkung der Limitierung der gleichzeitig aktiven Tests Rechnung getragen (vgl. Abschnitt 3.4.1), muss an dieser Stelle aber noch erweitert werden:

Bedingung 7: *Der Zeitraum zur Datenerfassung ist so weit zu strecken, dass ein Proband zu keinem Zeitpunkt einem übermäßig hohen Maß an Stress ausgesetzt ist.*

Darüber hinaus sollen die Probanden vor möglichen Konsequenzen durch ihre Teilnahme geschützt werden. Dieser Schutz ist dann gewährleistet, wenn die Identität der Person zu den erfassten Messergebnissen vertraulich behandelt wird. Um dem Bedürfnis nach einer vertraulichen Behandlung der Identität der Probanden nachzukommen, wird ein Pseudonymisierungs- und Anonymisierungskonzept erstellt, das die Wahrung der Vertraulichkeit der Identität der Probanden gewährleistet:

Bedingung 8: *Das Werkzeug muss die Identität der Probanden optimal schützen.*

Die Ergebnisse dieser Arbeit erlauben eine bessere Erfassung des IT-Sicherheitsbewusstseins von Mitarbeitern. Das gewählte Klinikszenario ermöglicht den Einsatz der Methode und dem zugehörigen Werkzeug zur artefaktbasierten Messung von IT-Sicherheitsbewusstsein in Kritischen Infrastrukturen. Damit sind die Ergebnisse dieses Experiments gesamtgesellschaftlich relevant. Die im Rahmen dieses gesamten Kapitels aufgestellten Bedingungen adressieren die identifizierten Risiken für die herausgearbeiteten Stakeholder. Werden Maßnahmen

ergriffen, diese Bedingungen nach dem Stand der Technik zu erfüllen, ist davon auszugehen, dass es sich um ein Experiment wohltätigen Charakters handelt.

4.2.4 „Justice“

Um das Prinzip der Gerechtigkeit anzuwenden, muss gezeigt werden, dass die Chance auf den Mehrwert und die Verteilung der Last unter allen Probanden vergleichbar ist. Dies betrifft insbesondere die Selektion als Proband und die Teilnahme an der Intervention.

Die Probandenselektion ist gerecht, da alle als Proband infrage kommenden Mitarbeiter auch selektiert werden. Allen Probanden soll die gleiche Chance gewährt werden, an der IT-sicherheitsbewusstseinserhöhenden Maßnahme teilzunehmen. Die Kosten für die Durchführung werden durch das ITS.APT-Projekt finanziert. Die Raummiete und die Kosten, die durch die Freistellung der Mitarbeiter entstehen, übernimmt der Anwendungspartner USKH selbst. Dadurch entstehen den Teilnehmern keine persönlichen Kosten. Eine Schulung (vgl. Abschnitt 2.3) lässt sich jedoch nur bedingt kostenneutral skalieren. Die zur Verfügung stehende Raumkapazität ist wie die Anzahl der möglichen Termine beschränkt:

Bedingung 9: *Jeder Proband soll die gleiche Chance erhalten, an der Schulung teilzunehmen.*

Aus diesem Grund wird die Teilnahme an der Maßnahme durch eine Einladung gesteuert. Diese wird mit der Nutzeraufklärung (vgl. Bedingung 4) zusammengefasst. Die Details zur Selektion für die Intervention sind in Abschnitt 6.1.4 dargestellt.

Unter den oben geschilderten Rahmenbedingungen wird angenommen, dass das Experiment gerecht ist.

4.2.5 „Respect for Law and Public Interest“

Im Rahmen des Forschungsprojekts „ITS.APT: IT-Security Awareness Penetration Testing“ (vgl. Abschnitt 1.2) entstand zur Gewährleistung der Gesetzmäßigkeit ei-

ne eingehende Bewertung des Vorhabens bezüglich der relevanten Rechtsgebiete Arbeitsrecht [Hey+16], Haftungsrecht [HO16] sowie Datenschutzrecht [JO16].

Die oben aufgelisteten juristischen Betrachtungen bilden die Grundlage für die im Rahmen des Projekts erarbeiten Handlungsempfehlungen [Bie+18]. Es sind konkrete Empfehlungen an Betreiber und Testanwender zum rechtssicheren Einsatz der im Rahmen dieser Arbeit entwickelten Methode. Wird den Handlungsempfehlungen Folge geleistet, kann davon ausgegangen werden, dass die Durchführung des Experiments gesetzmäßig ist. Aus diesem Grund werden diese Handlungsempfehlungen an dieser Stelle gezielt aufbereitet.

Zunächst wird festgestellt, dass die europäische Datenschutzgrundverordnung [GDP06] sowie das Bundesdatenschutzgesetz oder die Landesdatenschutzgesetze der jeweiligen Länder [BDS19, § 1 Anwendungsbereich des Gesetzes] Anwendung finden. Weiter wird ausgeführt, dass es sich bei den während der Durchführung der Studie erhobenen Daten um personenbezogene Daten der Beschäftigten handelt. Daraus ergibt sich, dass die verarbeiteten Daten zu minimieren und deren Erhebung weitestmöglich zu vermeiden ist. Die Daten dürfen nicht zum Zweck der individuellen Leistungskontrolle eingesetzt werden. Personenbezogene Daten Dritter dürfen weder durch das Werkzeug noch durch die Artefakte erfasst werden.

Diese Forderungen bekräftigen Bedingung 8 und die damit einhergehende Forderung nach einem Pseudonymisierungs- und Anonymisierungskonzept sowie die durch Bedingung 5 formulierte Isolation der Tests und damit der Probanden von den Nichtprobanden. Darüber hinaus ist es notwendig, die Datenminimierung und -sparsamkeit explizit festzuhalten und damit eine Rechtfertigung jedes erhobenen Datums zu erzwingen:

Bedingung 10: *Die in dem Verfahren zu verarbeitenden Daten sind zu minimieren.*

Mit der Anwendung der Datenschutznormen stellt sich auch die Frage der Rechtsgrundlage des Einsatzes. Der Einsatz der Messmethode, wie in Kapitel 3 beschrieben, ist auf Grundlage einer Betriebs- oder Dienstvereinbarung (je nachdem, ob es sich um ein privatwirtschaftliches Unternehmen oder eine öffentliche Einrichtung handelt) mit dem deutschen Recht konform. Diese kann der Betriebs- oder Personalrat für die Mitarbeiter schließen [Bie+18; Bet20]. Der

Abschluss einer Betriebs- oder Dienstvereinbarung wird sogar explizit empfohlen, um die Rechte der Personalvertretung zu wahren.

Darüber hinaus ist ein Eintrag in dem entsprechenden Verfahrensverzeichnis (vgl. [GDP06, Art. 30; BDS19, § 70]) anzulegen. Dieser wurde entsprechend angelegt. Ein Muster wurde im Rahmen des Projekts erarbeitet und veröffentlicht [Bie+18, Annex II].

Die Durchführung einer Datenschutzfolgeabschätzung ist im allgemeinen Fall zu empfehlen [Bie+18]. Der Bedarf nach dieser ergibt sich aus Artikel 35 Absatz 3 der DSGVO [GDP06]. Hier ergeben sich zwei mögliche Anknüpfungspunkte. Zum einen betrifft die Forderung nach dieser Abschätzung Daten, die als Grundlage einer Entscheidung dienen, welche die betreffende Person in erheblicher Weise beeinträchtigen kann [GDP06, Art. 35 Abs. 3a]. Durch die Forderung nach dem Schutz der Identität der Probanden (vgl. Bedingung 8) ist nicht davon auszugehen, dass dieser Fall eintritt. Der zweite Anknüpfungspunkt ist die Verarbeitung von Daten nach Artikel 9 Absatz 1 der DSGVO [GDP06, Art. 35 Abs. 3b]. Das Werkzeug verarbeitet Daten dieser Art nicht. Die Möglichkeit ist, je nach Ausgestaltung der zum Einsatz kommenden Artefakte, nicht ausgeschlossen. Entweder muss eine Datenschutzfolgeabschätzung durchgeführt werden oder die Artefakte dürfen keinen Rückschluss auf Daten der entsprechenden Klassen zulassen.

Bedingung 11: *Artefakte dürfen nicht so gestaltet sein, dass „Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung" [GDP06, Art. 9 Abs. 1] verarbeitet werden.*

Die Bedingung ist für die zum Einsatz kommenden Artefakte erfüllt (vgl. Anhang A). Von einer Datenschutzfolgeabschätzung wird aus diesem Grund abgesehen.

Haftungsrisikoreduktion

Aus den in der haftungsrechtlichen Risikoabschätzung [HO16] analysierten Risiken ergeben sich vier Bereiche, in denen Handlungsempfehlungen zur Reduktion eines Haftungsrisikos für den Testanwender und den Betreiber dokumentiert werden [Bie+18, II. Handlungen zur Haftungsreduzierung]. Folgend werden diesen entsprechende Bedingungen und Anforderungen entnommen.

Organisationspflichten im Betrieb Ein möglicher Schaden Dritter ist bereits durch die sorgfältige Organisation der Testdurchführung zu vermeiden. Der Testanwender ist vom Betreiber der Aufgabe angemessen auszuwählen, anzuleiten und zu überwachen. Eine schriftliche Protokollierung der Organisation wird empfohlen. Dies muss manuell durch den Testanwender erfolgen.

Definition geeigneter Testbereiche Einen KRITIS-Betreiber treffen insbesondere Pflichten gegenüber den Nutznießern der Diensterbringung. Deshalb muss die Empfehlung wie folgt lauten: „Der Test ist daher so zu gestalten, dass er nur solche Unternehmensbereiche berührt, die strukturell nicht essentiell für den Betrieb sind“ [Bie+18, S. 68]. Diese Definition ist technisch abzusichern. Die Empfehlung greift Bedingung 5 auf und verlangt darüber hinaus deren technische Absicherung:

Anforderung 5: *Das Werkzeug muss Probanden innerhalb der Infrastruktur identifizieren und darf Artefakte nur diesen präsentieren.*

Durchführung von Testläufen Die Durchführung von Testläufen wird explizit empfohlen [Bie+18]. Dabei wird die Installation auf korrekte Funktionalität hin geprüft. Diese Prüfung soll sowohl das zum Einsatz kommende Werkzeug als auch alle Artefakte umfassen. Diese Forderung greift Bedingung 6 (Seite 58) wieder auf und fügt einen konkreten Implementierungsansatz hinzu. Da Bedingung 6 dieses Vorgehen bereits impliziert, lässt sich keine neue Anforderung ableiten.

Überwachung der Tests an sich, keine unzulässige Datenverarbeitung Die Durchführung der Tests ist zu überwachen und bei Auffälligkeiten im Ablauf entsprechend zu reagieren. Das Werkzeug kann zur Unterstützung der Erfüllung dieser Forderung die Durchführung testrelevanter Aktionen dokumentieren:

Anforderung 6: *Alle Vorgänge durch das Werkzeug in der IT-Infrastruktur müssen protokolliert werden.*

4.3 Fazit

Im Rahmen dieses Kapitels wurden die ethischen und rechtlichen Rahmenbedingungen erfasst und aufbereitet. Dazu wurden zunächst verwandte Arbeiten herangezogen. Diese konnten jedoch nicht vollständig auf das geplante Vorhaben übertragen werden. Das Vorhaben wurde dann, durch die Prinzipien des Menlo-Reports geführt, ethisch bewertet. Entsprechende Abwägungen wurden in Anforderungen und Bedingungen festgehalten. Diese wurden im Anschluss, gestützt auf die im Rahmen des Forschungsprojekts ITS.APT (vgl. Abschnitt 1.2) erarbeiteten Handlungsanweisungen, um die rechtlichen Aspekte angereichert.

Es wurde ein Schema zur Probandenaufnahme in IT-Sicherheitsbewusstseinsmessungen von Angestellten entworfen, das auf der stellvertretenden Einwilligung durch den Betriebs- oder Personalrat beruht. Eine entsprechende Dienstvereinbarung wurde mit dem Gesamtpersonalrat der UKSH geschlossen. Ein Muster wurde im Rahmen des Projekts erarbeitet und hängt den Handlungsempfehlungen an [Bie+18].

Einen alternativen Entwurf stellen Resnik & Finn erst 2018, dem Projekt nachfolgend, in ihrer Arbeit „*Ethics and Phishing Experiments*“ [RF18], vor. Sie schlagen vor, die Mitarbeiter als Probanden über die Forschung vor der Datenerhebung lediglich generisch zu informieren. Eine vollumfängliche Information soll nach der Datenerhebung erfolgen. Vorab soll den Probanden nur bekannt gemacht werden, dass Forschung zu IT-Sicherheitsverhalten betrieben wird und sie das Recht haben, die Teilnahme zu verweigern. Dabei gehen sie jedoch nicht auf entsprechende juristische Rahmenbedingungen ein, die eine Beteiligung des Betriebs- oder Personalrats oder eine umfangreichere Information der Probanden erforderlich machen. Damit ist der Vorschlag von Resnik &

Finn nicht direkt umsetzbar. Wird er durch die Beteiligung des Betriebs- oder Personalrats angepasst, kann auf die individuelle Ansprache der Probanden verzichtet werden und es ergibt sich das oben hergeleitete Schema.

Durch die Aufarbeitung der ethischen und juristischen Anforderungen ist der Grundstein der Antwort auf die vierte Forschungsfrage nach der ethischen und rechtskonformen Umsetzung von IT-Sicherheitsbewusstseinsmessungen (vgl. Abschnitt 1.3) gelegt. Wie die entsprechenden Anforderungen und Bedingungen adressiert werden, ist den folgenden Kapiteln zu entnehmen. Kapitel 5 dokumentiert den Entwurf eines zu den ethischen und juristischen Rahmenbedingungen konformen Werkzeugs und bildet damit den zweiten und letzten Teil der Antwort auf die vierte Forschungsfrage.

5 Ein Werkzeug zur artefaktbasierten IT-Sicherheitsbewusstseinsmessung

Im Rahmen dieses Kapitels wird das entwickelte Werkzeug, das *IT-Security Awareness Penetration Testing Environment – ITS.APE*, zur artefaktbasierten IT-Sicherheitsbewusstseinsmessung vorgestellt. Damit wird die vierte Forschungsfrage vollständig beantwortet.

Die Beschreibung des Werkzeugs orientiert sich an dem Grundkonzept aus Kapitel 3. Analog lässt sich die Messung in drei Phasen unterteilen:

Aufbauphase Die *Aufbauphase* umfasst die *Artefaktkonstruktion* und deren Personalisierung (vgl. Anforderung 2 auf Seite 49).

Testdurchführung Während der *Testdurchführung* findet die kontrollierte *Artefaktpräsentation* (vgl. Anforderung 1 auf Seite 49) und das Aufzeichnen von Nutzerreaktionen (vgl. Anforderung 3 auf Seite 49) statt.

Nachbereitung Der Abbau des Testaufbaus und die Auswertung der erfassten Daten (vgl. Anforderung 4 auf Seite 49) finden in der *Nachbereitungsphase* statt.

Das erste Unterkapitel greift die in Abschnitt 3.2 zusammengetragenen Artefakte erneut auf und arbeitet die benötigten Fähigkeiten eines Werkzeugs zur Unterstützung dieser Artefakte heraus. Die zuvor herausgearbeiteten Bedingungen werden im Laufe dieses Kapitels entsprechend adressiert. Die folgenden Unterkapitel reflektieren die drei Phasen der Messung. Das Kapitel schließt mit einem Fazit.

5.1 Artefakttypen

Alle Artefakte, insbesondere die in Kapitel 3 beschriebenen, lassen sich in vier Klassen unterteilen. Je nach *Artefakttyp* haben diese unterschiedliche Anforderungen. Um diesen gerecht werden zu können, unterstützt ITS.APE eine Menge an Konfigurationen als Schnittstelle zu den Artefakten. Die Einteilung dieser Klassen ist in Abbildung 9 dargestellt.

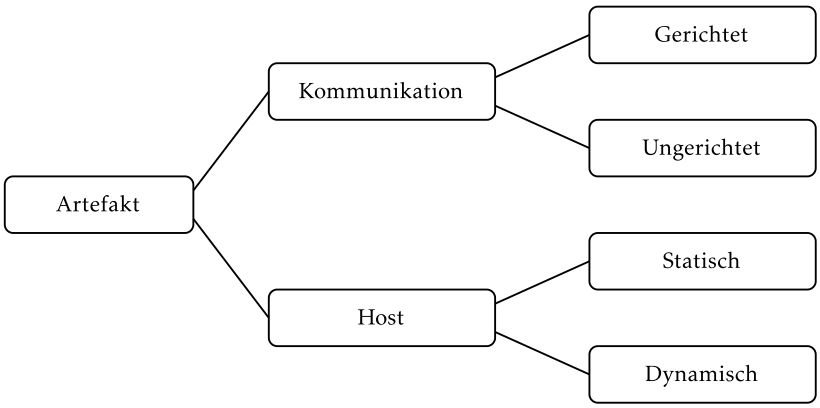


Abbildung 9: *Klassen von Artefakten.*

Artefakte werden in *kommunikationsbasiert* oder *hostbasiert* unterschieden. Kommunikationsbasierte Artefakte werden zur Präsentation von außen in einen bestehenden Kommunikationskanal eingegeben. Hostbasierte Artefakte existieren im Nutzerkontext des lokalen Betriebssystems [Mic18, Security boundries]. Kommunikationsbasierte Artefakte teilen sich in *gerichtete* und *ungerichtete* Artefakte auf. Das teilende Kriterium ist die Adressierbarkeit des Artefakts. Ein Beispiel für gerichtete kommunikationsbasierte Artefakte sind E-Mails, wie sie auch in der Studie zum Einsatz kommen: IT Ticket (vgl. Anhang A S. 124) oder Targo Bank (vgl. Anhang A S. 122). Eine E-Mail kann mithilfe der E-Mail-Adresse von außen direkt an einen Probanden adressiert werden. Ungerichtete kommunikationsbasierte Artefakte sind Artefakte, die der Nutzer wahrnehmen kann, obwohl die Kommunikation nicht gerichtet ist. Dies ist zum Beispiel der Fall, wenn ein Nutzer ein Artefakt beim Webbrowsing wahrnimmt. Als Repräsentanten kommen hier Artefakte wie das im Experiment eingesetzte Defacing (vgl. Anhang A S. 128) oder ein Online-Miner (vgl. Abschnitt 3.2.2) infrage. Um die in Bedingung 5 (Seite 57) geforderte Nichtbeeinträchtigung der

Nichtprobanden zu gewährleisten, müssen Probanden vor der Präsentation des Artefakts in dem entsprechenden Kommunikationskanal identifiziert werden.

Hostbasierte Artefakte teilen sich in *dynamische* und *statische* Artefakte auf. Für alle hostbasierten Artefakte gilt, dass das Artefakt in die Sitzung der Probanden injiziert werden muss. Statische Artefakte haben ohne die Interaktion des Probanden keinen Laufzeitkontext. Ein Beispiel für Artefakte dieser Klasse sind die durch eine Ransomware verschlüsselten Dateien (vgl. Abschnitt 3.2.2) oder Dateien unbekannter Herkunft wie zum Beispiel eine Selbstlöschende Datei (vgl. Anhang A S. 119) oder das Word Macro (vgl. Anhang A S. 133). Dynamische hostbasierte Artefakte müssen zudem noch aktiviert, d. h. als Prozess im Betriebssystem gestartet, werden. Beispiele für diese Artefakte finden sich insbesondere in Nutzerdialogen wie Java Update (vgl. Anhang A S. 132) und Login-Fenster (vgl. Anhang A S. 127).

Prinzipiell muss auch der Fall behandelt werden, dass sich ein Proband während der Testdurchführung abmeldet und ein Nichtproband sich anmeldet. In diesem Fall muss das Artefakt deaktiviert werden können. Werden alle diese Artefakttypen durch ITS.APE unterstützt, ist davon auszugehen, dass genug Varianz in den Artefakten erzeugt werden kann, um Bedingung 1 (Seite 48), der Präsentation von mehreren unterschiedlichen Artefakten, nachzukommen.

5.2 Aufbauphase

Die Aufbauphase von ITS.APE beinhaltet die Inbetriebnahme aller Softwarekomponenten, die Konfiguration der Testdurchführung und deren Ablaufplanung.

5.2.1 Inbetriebnahme

ITS.APT besteht aus zwei Hauptkomponenten, einer *Serverkomponente* (folgend nur *APE*) und einer *Clientkomponente* (folgend *Client*). APE ist eine zentrale Sammlung von Diensten und Werkzeugen, die für die Testdurchführung benötigt werden. Sie wird pro Installation in einer IT-Infrastruktur nur einmalig instanziiert. Diese Instanz muss von jedem Bildschirmarbeitsplatz, den die Probanden während des regulären Betriebs bedienen, folgend *Probandenrechner*,

aus über das Netzwerk direkt erreichbar sein. Der Testanwender steuert APE über eine Schnittstelle für die Kommandozeile.

Der Client wird auf jedem Probandenrechner installiert. Er wird als Windows Installer Paket [SSS18] zu Verfügung gestellt und kann über gängige Lösungen zur Softwareverteilung auf den entsprechenden Rechnern installiert werden, ohne dass die Probanden davon Notiz nehmen. Sind die Bildschirmarbeitsplätze personalisiert, empfiehlt es sich, den Client auch nur auf den Probandenrechnern zu installieren, um schon in diesem Schritt Bedingung 5 (Seite 57) nachzukommen. Die Funktionalität des Clients ist in Abschnitt 5.3.1 beschrieben.

Um alle Artefakte auf ihre korrekte Funktion zu prüfen, erfolgt die Installation des Clients in zwei Stufen. In der ersten Stufe wird der Client nur auf einem durch den Testanwender kontrollierten Bildschirmarbeitsplatz installiert. Dieser wird dann genutzt, um alle Artefakte, die in der bevorstehenden Testdurchführung zum Einsatz kommen sollen, auf ihre ordnungsgemäße Funktion hin zu testen. Dazu wird die gesamte Konfiguration und Testdurchführung mit dem Bildschirmarbeitsplatz des Testanwenders durchgeführt. Erst wenn alle Artefakte und APE getestet wurden, darf der Client in der zweiten Stufe an die Probandenrechner verteilt werden. Die Installation auf dem Probandenrechner kann mit einem Artefakt geprüft werden, das keinen durch die Probanden wahrnehmbaren Effekt erzeugt. Dieses Vorgehen erfüllt auch die Forderung nach einem Funktionstest des ausgebrachten Werkzeugs wie in Abschnitt 4.2.5 gefordert und beendet damit die Inbetriebnahme (vgl. Bedingung 6).

5.2.2 Konfiguration

Die Konfiguration der Testdurchführung beinhaltet die Übergabe der Probandendaten an APE, die Konfiguration der Artefakte, die Ablaufplanung und die technische Vorbereitung der Artefakte zur Präsentation zum Abschluss dieser Phase. Dieser Ablauf ist in Abbildung 10 auf der nächsten Seite dargestellt. Er beginnt nach der Inbetriebnahme aller Softwarekomponenten. Zunächst wird der Gesamttestzeitraum durch den Testanwender konfiguriert. Alle Tests finden innerhalb dieses Zeitraums statt. Die folgenden Abschnitte stellen die diesem Schritt folgenden Aktionen vor.

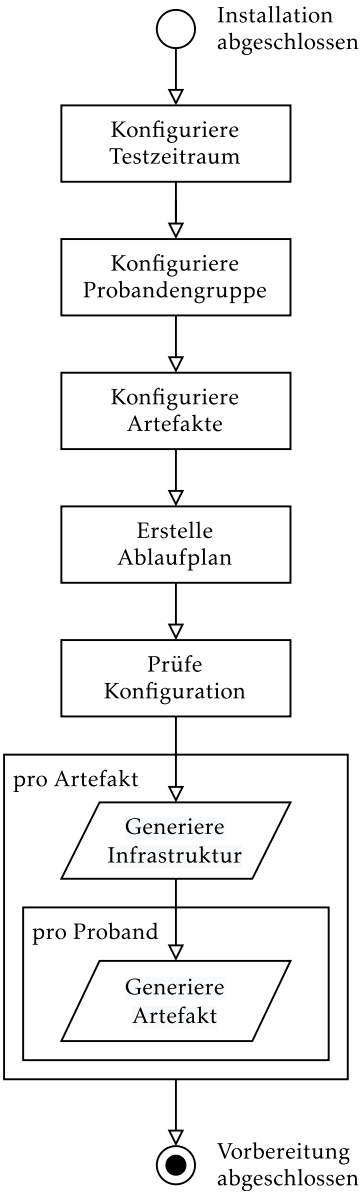


Abbildung 10: *Ablaufdiagramm zur Vorbereitung der Testdurchführung während der Aufbauphase. Rechtecke mit zentrierter Beschriftung geben Tätigkeiten an. Umfassende Rechtecke stellen Schleifen dar. Parallelogramme stellen Tätigkeiten mit Ausgaben dar.*

Probandenkonfiguration

Nachdem der Gesamttestzeitraum konfiguriert wurde, müssen die Probandendaten an APE übergeben werden. Abbildung 10 auf der vorherigen Seite benennt diesen Schritt mit *Konfiguriere Probandengruppe*. Jede Person, die nicht in diesem Schritt als Proband konfiguriert wird, ist Nichtproband.

Die Probanden werden an APE in Form einer CSV-formatierten Datei [Sha05] übergeben. Jede Zeile entspricht einem Probanden. Mindestens müssen die Daten *Vorname*, *Nachname*, *UserID* und *E-Mail* angegeben sein. Der Name des Probanden ist erforderlich zur Personalisierung von Artefakten. Die E-Mail-Adresse wird genutzt, um Phishing-E-Mails verschicken zu können. Die *UserID* ist das Erkennungstoken innerhalb der Domäne, typischerweise der Nutzername.

APE akzeptiert darüber hinaus beliebige weitere Felder. Damit lassen sich beispielsweise Untergruppen beschriften. Denkbar sind Abteilungsbezeichnungen oder Kennzeichen verschiedener Interventionen. Diese Beschriftungen können zur Auswertung der Ergebnisse wieder herangezogen werden.

Um die Identität der Probanden zu schützen, referenziert APE die Probanden wie in Abbildung 11 gezeigt. Die `id` kann durch eine entsprechende Spalte in der Gruppenbeschreibung frei vergeben werden. Wird die Gruppenbeschreibung wieder gelöscht, verlieren die durch APE erfassten Daten ihren Personenbezug. Um die Rekonstruktion der Identität zu erschweren, kann die `id` zufällig vergeben werden. Dies kann beispielsweise durch die Vertauschung der Zeilenreihenfolge innerhalb der Datei vor der Vergabe der `id` erfolgen. Damit ist Bedingung 8 (Seite 59) erfüllt.

Schema:	**Pseudonym:**
probandengruppe/id	`gruppe1.csv/73`

Abbildung 11: *Pseudonyme der Probanden.*

Artefaktkonfiguration

Die zum Einsatz kommenden Artefakte müssen für die Konstruktion und Präsentation konfiguriert werden. Abbildung 10 auf der vorherigen Seite benennt diesen Schritt mit *Konfiguriere Artefakte*. Artefakte sind in Form von *Rezepten*

als Beschreibungen abgelegt. Ein Rezept ist eine YAML-Konfiguration [BEd09], die eine Sammlung aus Anweisungen und Werten enthält. Das Rezept bestimmt alle artefaktbezogenen Konfigurationen. Der schematische Aufbau eines Rezepts ist in Abbildung 12 dargestellt. Ein Rezept enthält insbesondere die Anweisung zur Generierung (*ArtifactScript*) und Steuerung (*ArmScript*, *DisarmScript*, *DeployScript* und *UndeployScript*) von dem spezifischen Artefakt. Ein Rezept beinhaltet eine beliebige Anzahl von *Infrastrukturelementen* und *Trackfiltern*. Diese werden in den entsprechenden Phasenabschnitten der Testdurchführung detailliert beschrieben.

In dieser Phase werden die Rezepte zur Testung ausgewählt und parametrisiert. Insbesondere muss die Dauer der Testung festgesetzt werden. Sie ist für jeden Probanden, der einen Test auf Basis dieses Rezepts durchläuft, gleich. Darüber hinaus kann ein Rezept weitere beliebige Parameter unterstützen. Hier kann zum Beispiel die Absenderadresse für E-Mail-Artefakte angeboten werden.

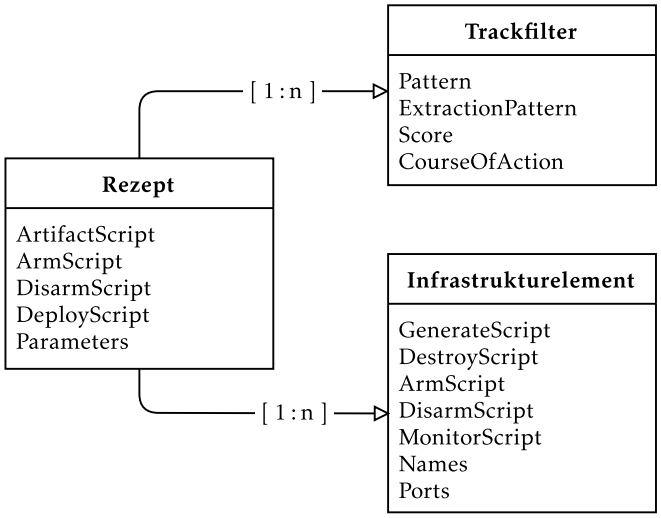


Abbildung 12: *Schematische Darstellung eines Rezepts. Entitäten eines Rezepts und die Kardinalität ihrer Beziehung.*

5.2.3 Ablaufplanung

Nach der Festlegung der Probandengruppe und der Artefakte in Form von Rezepten wird eine Ablaufplanung generiert. Abbildung 10 auf Seite 71 bezeichnet diesen Schritt mit *Erstelle Ablaufplan*. Dazu wird für jeden Einzeltest,

der sich aus der Kombination von Proband und Rezept ergibt, im Rahmen des Gesamttestzeitraums durch APE ein individueller Testzeitraum bestimmt. Die Laufzeit der Tests hat der Testanwender bei der Konfiguration der Artefakte bestimmt. APE verteilt diese nun auf den Gesamttestzeitraum. Je größer der Gesamtmesszeitraum ist, desto größer sind die Abstände zwischen den individuellen Testzeiträumen. Die Bestimmung der Testzeiträume muss die üblichen Arbeitszeiten der Mitarbeiter sowie Wochenenden und Feiertage beachten. Nur so kann von der Anwesenheit der Probanden während der Testzeiträume ausgegangen werden. APE bietet für diese Einschränkung entsprechende Konfigurationsoptionen und errechnet einen Zeitplan für die Testdurchführung.

Die Anordnung der Testzeiträume kann linear oder randomisiert erfolgen. Ist die Anordnung linear, werden die Tests nacheinander, Rezept für Rezept, abgespielt. Werden die Testdurchführungen randomisiert, ergibt sich für jeden Probanden ein individueller Ablaufplan und mehrere Rezepte sind gleichzeitig aktiv. In jedem Fall ist zu jedem Zeitpunkt für einen Probanden nur ein Test aktiv.

Die lineare Anordnung der Testdurchführung hat den Vorteil, dass beim Auftreten unvorhersehbarer Ereignisse die erfassten Daten erhalten bleiben. Jedoch hat es auch den Nachteil, dass die Probanden gleichzeitig mit dem gleichen Artefakt konfrontiert werden. Dies kann dazu führen, dass sich Probanden über die Artefakte austauschen (vgl. [Kum+08, 4.5 Observations]). Bei der randomisierten Testanordnung wird diesem Effekt entgegengewirkt. Probanden werden so zu unterschiedlichen Zeiten unterschiedliche Artefakte präsentiert. Jedoch hat dieses Vorgehen auch Nachteile. Der Effekt schwindet, wenn der Gesamttestzeitraum eng gewählt wird. Die einzelnen Tests können in diesem Fall nur mit kleinem Zeitversatz ausgeführt werden. Die Menge der Probanden, die zu jedem Zeitpunkt einen Test mit dem gleichen Artefakt durchlaufen, steigt bis:

$$\frac{|\text{Probanden}|}{|\text{Artefakte}|} \tag{5.1}$$

Muss die gesamte Testdurchführung pausiert werden, muss beim Fortsetzen der Testdurchführung der Ablaufplan neu berechnet werden. Die Funktion, eine

pausierte Testdurchführung wieder aufzunehmen, wird von APE bisher nicht unterstützt.

5.2.4 Phasenabschluss

Zum Phasenabschluss wird der Ablaufplan freigegeben. Der durch APE errechnete Ablaufplan kann zu diesem Zeitpunkt allen involvierten Stellen wie dem Helpdesk, dem IT-Sicherheitsbeauftragten und dem Personalrat kommuniziert werden. Zu diesem Zweck kann der Ablaufplan depseudonymisiert exportiert werden. Abbildung 10 auf Seite 71 bezeichnet diesen Schritt mit *Prüfe Konfiguration*.

Ist die gesamte Konfiguration geprüft und freigegeben, werden alle Artefakte sowie die benötigte Infrastruktur durch APE konstruiert. Dafür greift APE auf das Rezept, wie in Abbildung 12 auf Seite 73 gezeigt, zurück. Für jedes gewählte Rezept wird für jedes benötigte Infrastrukturelement das *GenerateScript* aufgerufen. Dies erzeugt die unterstützende Infrastruktur.

Die Infrastruktur wird den Probanden *dynamisch* verfügbar gemacht. Ein Proband kann die Infrastruktur während des entsprechenden Testzeitraums über ihren DNS-Namen auf einem konfigurierten TCP- oder UDP-Port erreichen. Um dies zu ermöglichen, bietet APE einen eigenen Nameserver an und verwaltet die Firewall des Betriebssystems. Jedes Infrastrukturelement kann DNS-Namen und Ports, unter denen es verfügbar sein soll, frei wählen. Dem *GenerateScript* wird über eine Umgebungsvariable ein freier Port auf der Loopback-Schnittstelle zugewiesen. Netzwerkpakete der Probanden werden in dem spezifischen Testzeitraum von dem gewünschten Port auf diesen *internen* Port per „*Destination Network Address Translation*“ (oft kurz DNAT, vgl. auch [SH99, 4.1.2. Network Address Port Translation (NAPT)]) weitergeleitet. Weitere Informationen über das Routing des Netzwerkverkehrs auf die Infrastruktur sind in Abschnitt 5.3.2 zu finden.

Nebenläufig zur Konstruktion der Infrastruktur werden die Artefakte generiert. Hierzu wird das *ArtifactScript* eines Rezepts (siehe Abbildung 12 auf Seite 73) einmal pro Proband aufgerufen. Um die Personalisierung der Artefakte zu ermöglichen (vgl. Anforderung 2 auf Seite 49), werden die Daten des Probanden (vgl. Abschnitt 5.2.2) wie auch bei der Konstruktion der Infrastruktur als

Variablen in die Laufzeitumgebung verfügbar gemacht. So können beispielsweise personalisierte E-Mails oder Webseiten aus einer Vorlage generiert werden.

Um die konstruierten Artefakte auf der Infrastruktur zu platzieren, wird nach beiden oben geschilderten Prozessen das *DeployScript* aufgerufen. Nun wird die Infrastruktur an das interne Monitoring von APE übergeben. Dieses ruft periodisch die *MonitorScripts* der Infrastrukturelemente auf, um sicherzustellen, dass der Testaufbau einsatzbereit ist.

Alle Aktionen, die der Testanwender bis zu diesem Zeitpunkt durchgeführt hat, sowie alle Aktionen, die das Werkzeug steuert, insbesondere die Testdurchführung, werden pseudonymisiert geloggt (vgl. Anforderung 6 und Bedingungen 8 und 10).

5.3 Testdurchführung

Beginnt der Testzeitraum für einen Test, wird die Testdurchführung durch APE ausgelöst. Der schematische Aufbau der Testdurchführung ist in Abbildung 13 auf der nächsten Seite dargestellt. Der Status der Infrastrukturelemente wird kontinuierlich überwacht. Die Teilschritte der Testdurchführung werden nur ausgeführt, wenn die Infrastrukturelemente verfügbar sind. Sind die Infrastrukturelemente zu einem Zeitpunkt während der Testdurchführung nicht mehr verfügbar, wird die Testdurchführung aller Instanzen dieses Rezeptes beendet und die in Abschnitt 5.3.5 beschriebenen Aktionen werden durchgeführt. Der weitere Ablauf der Testdurchführung ist in den folgenden Abschnitten beschrieben.

5.3.1 Der Client

Ist der Client auf dem Probandenrechner installiert, stellt dieser einen lokalen Hintergrunddienst in Form eines Windows Services zur Verfügung. Der Client erfüllt drei Rollen:

1. Die Verfügbarkeit des Netzwerks prüfen und anhand des Ergebnisses die Testteilnahme verwalten.

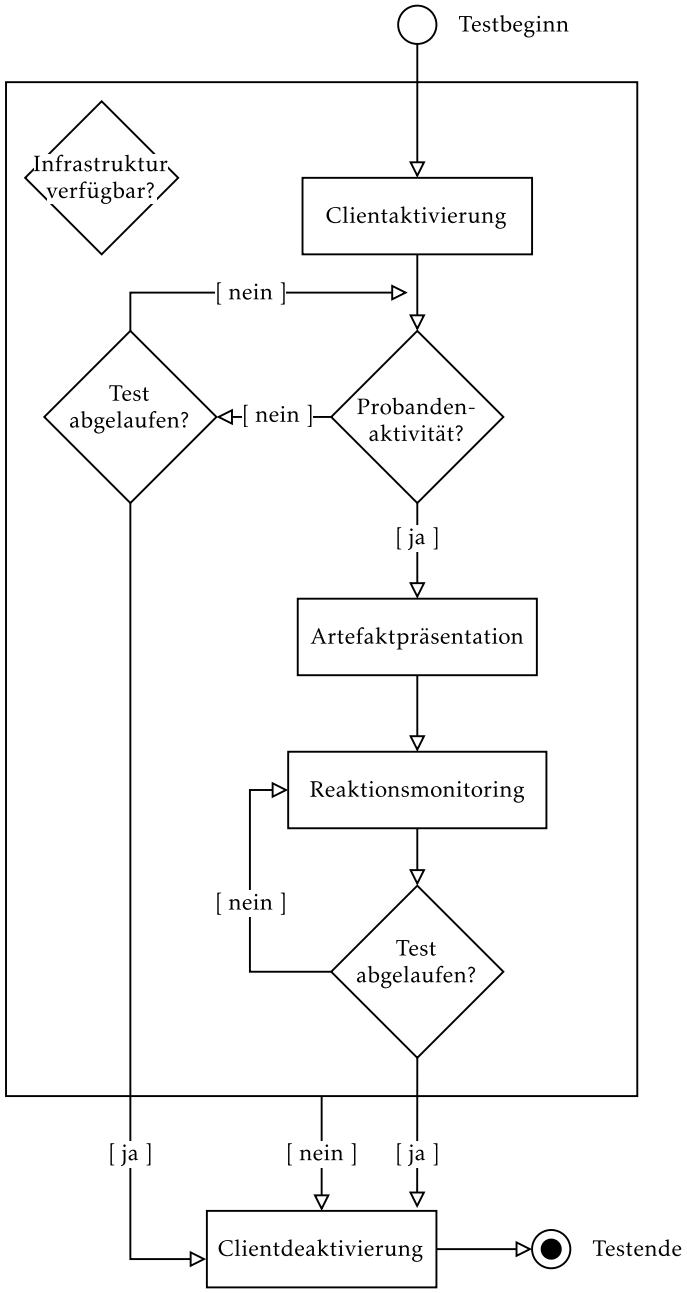


Abbildung 13: *Ablaufdiagramm der Datenerhebung während der Testdurchführung. Rechtecke mit zentrierter Beschriftung geben Tätigkeiten an. Rauten repräsentieren Entscheidungen. Mit Rauten markierte Rahmen stellen kontinuierlich geprüfte Entscheidungen wie die Prüfung der Infrastrukturerreichbarkeit dar.*

2. Die Anwesenheit des Probanden feststellen und diese an die Serverkomponente melden.
3. Die Präsentation der Artefakte unterstützen.

Zur Konnektivitätsprüfung versucht der Client in einem festgelegten Zeitintervall einen vorkonfigurierten Domänennamen aufzulösen. Der Standardwert ist aufgrund seiner hohen Beliebtheit `www.google.com`. Kann dieser aufgelöst werden, fragt der Client den Nameserver von APE an. Er fragt hier den Namen `its.apt` an. Lautet die Antwort des Nameservers `127.0.0.1`, ist an dem Rechner gerade ein Proband angemeldet. In diesem Fall wird der Nameserver, der durch APE bereitgestellt wird, als Nameserver des Probandenrechners konfiguriert. Damit ist der Client wie in Abbildung 13 auf der vorherigen Seite dargestellt *aktiviert*. In jedem anderen Fall wird die Konfiguration des Nameservers der Infrastruktur wieder hergestellt und der Client damit *deaktiviert*. Abbildung 14 illustriert diesen Ablauf.

Input : IP-Adresse von APE: *ape_ip*
Input : IP-Adresse von Infrastruktur Nameserver: *nameserver_ip*

```
1  while true do
2      wait(timer)
3      if current_dns.lookup("www.google.de").success? then
4          if ape_ip.lookup("its.apt") == "127.0.0.1" then
5              configure_dns(ape_ip)
6          else
7              configure_dns(nameserver_ip)
8          end
9      else
10         configure_dns(nameserver_ip)
11     end
12 end
```

Abbildung 14: *Pseudocode der Konnektivitätsprüfung des Clients.*

Die nächsten beiden Rollen erfüllt der Client dadurch, dass er in regelmäßigen Abständen APE nach *Aufträgen* für den aktuell angemeldeten Benutzer fragt. Ist kein Nutzer angemeldet, wird diese Anfrage nicht gesendet. Um diese Anfragen beantworten zu können, stellt APE einen Webserver bereit, der Arbeitsaufträge für die Clients bereithält. Diese Aufträge dienen der Artefaktpräsentation und der

Feststellung der Probandenaktivität. Sie werden in den folgenden Abschnitten erläutert.

5.3.2 Probandenaktivität und dynamisches Routing

Empfängt APE eine Anfrage nach Aufträgen, wird der Nutzername mit den Nutzernamen der hinterlegten Probanden verglichen, die sich gerade in einem Test befinden sollten. Ist das der Fall, wird die Zuordnung der IP der Anfrage mit dem Pseudonym des Probanden in einem geteilten, flüchtigen Speicher gelegt. Handelt es sich dabei nicht um einen in Testung befindlichen Probanden, werden die Daten verworfen. Wenn der Probandenrechner nicht in den üblichen Abständen mit dem Namen des Probanden nach Aufträgen fragt, geht APE davon aus, dass sich der Proband abgemeldet hat, und verwirft die Daten. Wie in Abbildung 13 auf Seite 77 dargestellt, wird ein Artefakt nur präsentiert, falls die Aktivität des Probanden festgestellt werden kann. Damit wird die Bedingung nach dem Tracking der Probanden erfüllt (vgl. Anforderung 5 auf Seite 63).

Die so erfassten Daten über die Probanden sind auch Grundlage des dynamischen Routings (vgl. Abschnitt 5.2.4). Aufgrund dieser Daten werden dynamisch die Firewallregeln erzeugt, die das NAT für Infrastrukturelemente implementiert und die Antworten des Nameservers gesteuert. Anhand der IP-Adresse der jeweiligen Anfrage lässt sich so feststellen, zu welchem Probanden diese Anfrage gehört. Da für einen Probanden zu jedem Zeitpunkt nur ein Test aktiv sein kann, ergibt sich aus der IP und dem Zeitpunkt der Anfrage eine Verbindung zu den Infrastrukturelementen des für den Probanden gerade aktiven Rezepts. Damit können für zwei Probanden zwei unterschiedliche Artefakte unter demselben Domänennamen und Port bereitgestellt werden.

5.3.3 Artefaktpräsentation

Die Artefaktpräsentation wird über das *ArmScript* des Rezepts und die Skripte der zugehörigen Infrastrukturelemente gesteuert (vgl. Abbildung 12 auf Seite 73). Während der Versand von E-Mails an dieser Stelle trivial zu implementieren ist, wird die Präsentation hostbasierter Artefakte durch den Client unterstützt. Dieser fragt APE nach zu bearbeitenden Aufträgen.

Der copy-Auftrag wird genutzt, um statische hostbasierte Artefakte vom Webserver herunterzuladen und zu platzieren. Es existieren drei unterschiedliche Typen von Aufträgen: *copy*, *remove* und *execute*. Aufträge des Typs execute werden genutzt, um dynamische hostbasierte Artefakte auf den Probandenrechner zu laden und in der Sitzung des Nutzers auszuführen. Remove-Aufträge dienen der Entfernung von Artefakten zum Testende.

Wird das Artefakt einem Nutzer präsentiert, wird dieses Ereignis festgehalten. Es ist denkbar, dass ein Proband im Verlauf des Testzeitraums nicht an seinem Rechner angemeldet ist. Ein Proband kann durch Krankheit, Urlaub, Dienstreisen oder dem Ausscheiden aus dem Unternehmen nicht an der Messung oder Teilen dieser Messung teilnehmen. Auch ist es möglich, dass der Proband von dem Zeitpunkt der Verteilung der Clients und der ersten Artefaktpräsentation wegen eines Defekts den Bildschirmarbeitsplatz wechseln muss. Nur durch die Aufzeichnung dieses Ereignisses lässt sich feststellen, an welchen Tests ein Proband während einer Messung tatsächlich teilgenommen hat.

5.3.4 Reaktionsmonitoring

Für das Reaktionsmonitoring (vgl. Anforderung 3 auf Seite 49) stellt APE einen eigenen Dienst bereit. Dieser akzeptiert Netzwerkverbindungen und nimmt übermittelte Logdaten entgegen. Diese Logdaten werden mithilfe der *Trackfilter* der Rezepte (vgl. Abbildung 12 auf Seite 73) geparst. Das Ziel ist das in Abbildung 7 auf Seite 44 gezeigte Datenformat aus Abschnitt 3.4.1.

Jeder Trackfilter besteht aus zwei regulären Ausdrücken, dem *Pattern* und dem *ExtractionPattern*, einer numerischen Angabe *Score* (deutsch: Punktzahl) und einer *CourseOfAction* (deutsch: Handlungsoption). Das *Pattern* wird während der Verarbeitung genutzt, um zu bestimmen, welchem Trackfilter der aktiven Rezepte, und damit zu welchem Test, ein Logdatum zuzuordnen ist. Ist dieser bestimmt, wird das *ExtractionPattern* des Trackfilters genutzt, um einen Identifikator für die Identität des Probanden und einen Zeitstempel zu extrahieren. Dafür werden *named captures* [Com16] genutzt. Zur Ermittlung der Identität kommen die Bezeichner der Kopfzeile der entsprechenden Gruppenbeschreibung (siehe Abschnitt 5.2.2) oder eine IP-Adresse (siehe Abschnitt 5.3.2) infrage. Kann kein Zeitstempel extrahiert werden, wird der Wert auf den Empfangszeitpunkt des Datums gesetzt. Kann eine Verbindung

hergestellt werden, wird die entsprechende Handlungsoption aus dem Trackfilter und ihre numerische Repräsentation mit abgespeichert. Wird kein sich in diesem Test befindlicher Proband identifiziert, werden die Daten gemäß Bedingung 10 (Seite 61) verworfen. Abbildung 15 illustriert diesen Vorgang.

```
Input : Logdata: logs
1  foreach active recipe: r do
2      foreach r.trackfilters: tf do
3          if log.match(tf.pattern) then
4              matches = log.match(tf.extractionPattern)
5              timestamp = matches[timestamp] : Date.now()
6              foreach Proband: p do
7                  if p ∈ matches and p.active(r, timestamp)? then
8                      store(p, r, timestamp, tf.courseOfAction, tf.score)
9                  end
10             end
11         end
12     end
13 end
```

Abbildung 15: *Pseudocode des Parsing beim Reaktionsmonitoring.*

Um die Logdaten an diesen Dienst übergeben zu können, stellt APE ein Programm zur Verfügung, das in der Lage ist, die Logeinträge einer beliebigen Anwendung an APE zu übergeben. Dieses Programm wird auch während der Infrastrukturkonstruktion, wie in Abschnitt 5.2.4 beschrieben, in der Laufzeitumgebung zur Verfügung gestellt.

5.3.5 Testende

Ist die Infrastruktur für die Testdurchführung nicht mehr erreichbar oder ist das Zeitintervall für den Test abgelaufen, ist das Testende erreicht (vgl. Abbildung 13 auf Seite 77). Wird das Testende durch den Ablauf des Zeitintervalls ausgelöst, wird die IP-Adresse-Probanden-Beziehung aus dem Speicher entfernt. Damit ist es Probanden nicht mehr möglich, auf die Infrastruktur zuzugreifen, und die entsprechenden Clients deaktivieren sich automatisch. Darüber hinaus wird das *DisarmScript* des Rezepts aufgerufen (vgl. Abbildung 12 auf Seite 73).

So können noch ausgebrachte Artefakte von den Probandenrechnern entfernt werden (vgl. Abschnitt 5.3.3). Ist der letzte Test eines Rezepts abgelaufen oder ist die Infrastruktur nicht mehr verfügbar, werden die durch die Infrastruktur reservierten Namen aus dem Nameserver und die entsprechenden Firewallregeln (siehe Abschnitt 5.2.4 auf Seite 75) entfernt.

5.4 Nachbereitungsphase

Die erfassten Verhaltensdaten der Probanden können nun pseudonymisiert aus APE exportiert werden. Ein nichtpseudonymisierter Export ist nicht vorgesehen, siehe dazu Bedingung 8 (Seite 59). Wird ein weiteres Feld der Probandengruppe exportiert (vgl. Abschnitt 5.2.2), setzt APE k-Anonymität [Swe02] durch. Dabei ist k durch den Testanwender zu spezifizieren, der Standardwert ist 3. Wird ein Studiendesign verfolgt, in dem die Messung für sich alleine steht (vgl. [SCC02, Designs That Use a Control Group But No Pretest]), kann die Gruppenbeschreibung entfernt werden. Folgen auf die Messung noch eine Intervention und eine weitere Messung, müssen diese Probanden dieser Maßnahmen mit den Probanden der abgeschlossenen Messung verkettet werden. Zu diesem Zweck bietet APE eine Schnittstelle an, um eine Gruppenbeschreibung *abhängig* zu einer vorher pseudonymisierten Gruppenbeschreibung zu pseudonymisieren. Die durch APE für Probanden in neuen Gruppenbeschreibungen generierten *ids* (vgl. Abbildung 11 auf Seite 72) werden übertragen, falls ein Proband reidentifiziert werden kann. Probanden können anhand der in Abschnitt 5.2.2 beschriebenen Mindestkonfiguration reidentifiziert werden. Bereits vergebene *ids* werden nicht erneut vergeben, falls kein Proband reidentifiziert werden kann. Wird die Pseudonymisierung unabhängig von einer weiteren Gruppenbeschreibung durchgeführt, werden die *ids* mehrfach vergeben. Abbildung 16 zeigt ein Beispiel.

Für die Auswertung der Daten stellt ITS.APE ein entsprechendes Softwarepaket zur Verfügung, das Analysen, wie sie in Abschnitt 6.3 dargestellt werden, unterstützt (vgl. Anforderung 4 auf Seite 49). Sind alle Tests durchgeführt, können die Clients von den Probandenrechnern über die entsprechenden Betriebssystemfunktionen entfernt werden. Sind die Reaktionsdaten und die

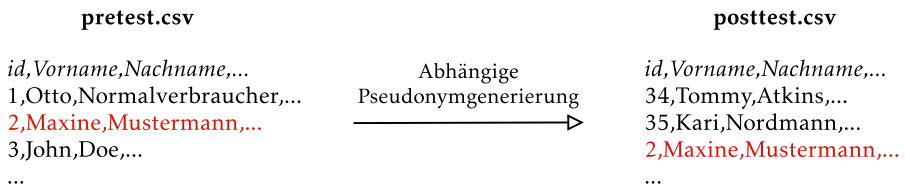


Abbildung 16: *Generierung verkettbarer Pseudonyme für voneinander abhängige Messungen. Der Proband „Maxine Mustermann“ wurde bei der Pseudonymisierung von* `posttest.csv` *auf Basis von* `pretest.csv` *reidentifiziert, ihm wird die gleiche* id *zugewiesen. Die anderen in* `posttest.csv` *beschriebenen Probanden wurden nicht reidentifiziert. Ihnen werden* ids *zugewiesen, die nicht in* `pretest.csv` *zum Einsatz kamen.*

Artefaktkonfiguration (vgl. Abschnitt 5.2.2) exportiert, kann der Server aus der Infrastruktur entfernt werden.

5.5 Fazit

Auf Basis des in Kapitel 3 dargestellten Konzepts wurde ein Werkzeug entworfen, das zur artefaktbasierten Messung von IT-Sicherheitsbewusstsein in der Lage ist. Die Darstellung in diesem Kapitel ist auf die Schnittstellen fokussiert, die APE bietet, und führt durch die drei Phasen der Messung: Aufbauphase, Testdurchführung und Nachbereitung.

Bei der Konzeption des Werkzeuges wurden alle identifizierten, technisch umsetzbaren Bedingungen beachtet. Tabelle 3 fasst die Kernaspekte zusammen. Die Bedingungen, die sich an die Organisation einer Studie richten, müssen beim Einsatz des Werkzeugs entsprechend gewürdigt werden. Die Erprobung des Werkzeugs wird im Folgekapitel beschrieben.

Die Darstellungen in diesem Kapitel liefern die zweite Teilantwort auf die Frage nach der ethischen und rechtskonformen Umsetzung artefaktbasierter IT-Sicherheitsbewusstseinsmessungen (vgl. Forschungsfrage 4). In Verbindung mit der Aufbereitung der ethischen und juristischen Rahmenbedingungen in Kapitel 4 ist durch den Entwurf eines Werkzeugs, das die artefaktbasierte Messung von IT-Sicherheitsbewusstsein implementiert, die Frage vollständig

Tabelle 3: *Übersicht der durch das Werkzeug berücksichtigten Bedingungen.*

Bedingung	Zusammenfassung
Bedingung 1: Ein Proband muss mehrere Tests mit unterschiedlichen Artefakten durchlaufen.	Das Werkzeug unterstützt alle möglichen Artefakttypen und begünstigt damit die erforderliche Varianz der Artefakte (vgl. S. 69).
Bedingung 5: Durch den Einsatz des Werkzeugs im Rahmen des Experiments darf kein Nichtproband beeinträchtigt werden, der nicht vorher explizit dem Vorhaben zugestimmt hat.	Der Client steuert die Ausbringung der Artefakte. Der Test ist nur auf den Rechnern aktiv, auf den ein Client installiert ist. (vgl. S. 70). Probanden werden auf diesen Bildschirmarbeitsplätzen anhand ihres Accountnamens identifiziert (vgl. S. 69).
Bedingung 6: Die Wahrscheinlichkeit für eine Beeinträchtigung der IT-Infrastruktur durch den Einsatz des Werkzeugs ist zu minimieren.	In Abschnitt 5.2.1 werden konkrete Empfehlungen zur Inbetriebnahme des Werkzeugs ausgesprochen (vgl. S. 70).
Bedingung 8: Das Werkzeug muss die Identität der Probanden optimal schützen.	Das Werkzeug referenziert Probanden über Pseudonyme (vgl. 72). Der Export der erfassten Reaktionen ist ausschließlich in pseudonymer Form unterstützt, die Auswertung ist anonym (vgl. 72).
Bedingung 10: Die in dem Verfahren zu verarbeitenden Daten sind zu minimieren.	Ausschließlich Daten, die einem Test zugeordnet werden können, werden gespeichert (vgl. 81).

beantwortet. Die Erprobung des Werkzeuges im Rahmen eines Experiments verifiziert diese Antwort.

6 Erprobung von artefaktbasierter IT-Sicherheitsbewusstseinsmessung durch ein Feldexperiment

Zur Evaluation der erarbeiteten Antworten auf die Forschungsfragen wird ein Feldexperiment durchgeführt. Es wird geprüft, ob die artefaktbasierte Messung von IT-Sicherheitsbewusstsein valide und reliabel ist. In dem durchgeführten Feldexperiment werden Probanden mehrere Artefakte (vgl. Bedingung 1 auf Seite 48) präsentiert und die Reaktionen der Probanden anschließend ausgewertet. Dieses Feldexperiment findet bei dem Anwendungspartner des ITS.APT-Projekts, dem Universitätsklinikum Schleswig-Holstein (vgl. Abschnitt 1.2), statt.

Die erfassten Reaktionen werden im Rahmen einer explorativen Datenanalyse ausgewertet. Als Nachweis der Validität der Messung wird mittels einer Messung auf Basis des entwickelten Messkonzepts der Effekt einer bewährten Intervention auf das IT-Sicherheitsbewusstsein der Mitarbeiter demonstriert. Dazu wird die Effektstärke der Intervention auf das konstruierte Maß (vgl. Abschnitt 3.4) demonstriert. Um die Reliabilität dieser Messung nachzuweisen, wird die *interne Konsistenz* [DB16] der Messung untersucht.

In den folgenden Abschnitten wird zunächst der Versuchsplan vorgestellt, in dem die Konzeption der Studie dargestellt wird. Darauf folgend wird die Durchführung des Experiments dokumentiert. In Abschnitt 6.3 werden die erfassten Daten ausgewertet und die Ergebnisse des Experiments dargestellt. Das Kapitel schließt mit einem Fazit.

6.1 Versuchsplan

In diesem Kapitel wird die Planung der Studie beschrieben. Da sich die Tests aufgrund der zum Einsatz kommenden Artefakte unterscheiden, wird als Studiendesign ein Nur-Posttest-Aufbau mit Kontrollgruppe und einer unabhängigen Pretest-Stichprobe [SCC02, Seite 115ff] zugrunde gelegt. Die weiteren Abschnitte beschreiben die Planung in der chronologischen Ordnung der Durchführung von der Probandenselektion über die Planung des Pretest, der Nachbesprechung der Messphase sowie der Intervention bis zum Posttest.

6.1.1 Testumgebung

Im Rahmen der Erfassung der Testumgebung (vgl. Abschnitt 3.1) sind insbesondere folgende Punkte von Interesse: Handlungsoptionen, die den Nutzern zur Verfügung stehen, die Aufnahme der technischen Gegebenheiten für den Entwurf der Artefakte sowie organisatorische Rahmenbedingungen wie etwa der Testzeitraum.

Handlungsoptionen

Mitarbeiter des Anwendungspartners, bei dem die Messung durchgeführt werden soll, haben sehr eingeschränkte Rechte auf ihrem direkten Bildschirmarbeitsplatz. Benötigt der Mitarbeiter Unterstützung oder sind Aktionen durchzuführen, die eine höhere Berechtigung am System voraussetzen, ist der Mitarbeiter angehalten, Kontakt mit dem Helpdesk als erste Anlaufstelle aufzunehmen. Dies ist insbesondere dann der Fall, wenn durch den Mitarbeiter eine sicherheitsrelevante Situation erkannt wird. Damit muss auch diese Handlungsoption erfasst werden (vgl. Anforderung 3). Jeder Mitarbeiter im Testfeld ist über diese Handlungsaufforderung schriftlich, durch die zu seiner Einstellung ausgehändigten Betriebsvereinbarung, informiert worden.

Das Helpdesk bietet zwei mögliche Kontaktkanäle an: eine Hotline und den Kontakt per E-Mail (siehe Abbildung 17). Die Kommunikation per Telefon ist synchron. Ist kein Mitarbeiter in der Lage den Anruf anzunehmen, bleibt der Nutzer in der Warteschleife. Dieser Zustand hält so lange an, bis der Nutzer auflegt oder ein Mitarbeiter den Anruf doch noch annehmen kann. Legt der Nutzer auf, bleibt es ihm überlassen, zu einem späteren Zeitpunkt erneut

anzurufen. Kommt der Kontakt mit dem Nutzer zustande, wird ein *Ticket*, das Abbild einer Anfrage, vom Mitarbeiter im System erzeugt. Kontaktiert der Nutzer das Helpdesk per E-Mail, wird aus der E-Mail automatisiert ein solches Ticket erzeugt.

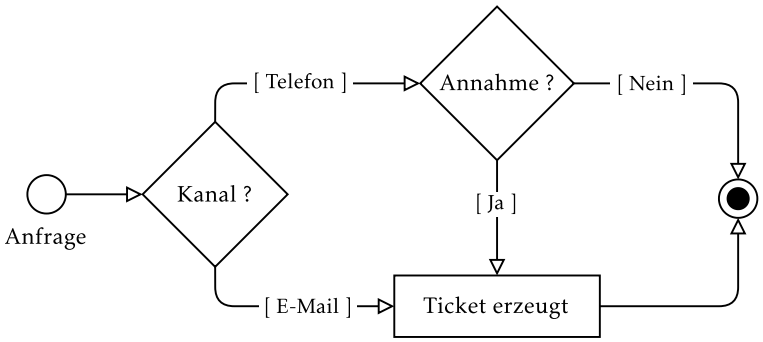


Abbildung 17: *Ablaufdiagramm für den Kontakt eines Nutzers zum Helpdesk.*

Soll dieser Kontakt als Handlungsoption aufgezeichnet werden, sind *beide* Kommunikationskanäle zu überwachen. Würden nur die erstellten Tickets aufgezeichnet werden, so würden nicht entgegengenommene Anrufe nicht mit erfasst werden. Dies würde keinen Rückschluss auf das IT-Sicherheitsbewusstsein der Probanden zulassen, sondern lediglich die Erreichbarkeit der Helpdesk-Mitarbeiter bewerten. Um diese Daten aufzunehmen, bietet APE eine Schnittstelle an, die sowohl in der Lage ist, einen Export aller Anrufe wie auch aller erstellten Tickets auszuwerten.

Damit wird dem Probanden unterstellt, dass dieser das Helpdesk erreichen wollte, um das Artefakt eines Tests zu melden. Ohne entsprechende Einwilligung der Probanden darf jedoch nicht auf den Inhalt der Kommunikation Bezug genommen werden (vgl. [Bie+18, 2. Telekommunikationsgesetz (TKG)]). Alle weiteren erfassten Handlungsoptionen ergeben sich durch die Interaktionsmöglichkeiten mit den Artefakten.

Technische Gegebenheiten

Die Konfiguration der Bildschirmarbeitsplätze ist zweigeteilt. Die Bildschirmarbeitsplätze auf den Krankenstationen werden von mehreren Mitarbeitern gemeinsam genutzt. Dies erfolgt so, dass ein Mitarbeiter sich einmalig lokal anmeldet.

Alle Mitarbeiter nutzen dann einen Applikationsserver, in dem die Anwendungen ausgeführt werden. Die Nutzung dieser Lösung setzt eine erneute Anmeldung an diesem Server voraus. Die Bildschirmarbeitsplätze in der Verwaltung sind hingegen personalisiert und die Mitarbeiter führen Applikationen auch lokal aus. Die unterschiedliche Konfiguration der Bildschirmarbeitsplätze erzeugt eine hohe Komplexität bei der Umsetzung von Anforderung 5 (Seite 63). Bei den Bildschirmarbeitsplätzen der Krankenstationen handelt es sich um patientennahe Systeme. Patienten sind als Nichtprobanden generell nach Bedingung 5 (Seite 57) zu schützen. Sie befinden sich darüber hinaus als Patienten in einer vulnerablen Position. Da es sich um den Ersteinsatz des neu konzeptionierten Werkzeugs handelt, wird der Test auf die Mitarbeiter der Verwaltung beschränkt.

Als Arbeitsstationen in der Verwaltung kommen während des Testzeitraums Rechner mit dem Betriebssystem Microsoft Windows 7 zum Einsatz. Abbildung 18 (Seite 89) zeigt die Softwarekonfiguration dieser Rechner. Artefakte müssen entsprechend entworfen werden und als solche erkennbar sein (vgl. Kapitel 3).

Testorganisation

Im Rahmen der Testorganisation wird der organisatorische Rahmen der Messung erfasst. Das Werkzeug wird anhand der in Abschnitt 5.2.1 ausgesprochenen Empfehlungen in Betrieb genommen (vgl. Bedingung 6).

Für die Messung wurde das Verfahren durch den zuständigen Datenschutzbeauftragten geprüft und ein entsprechender Eintrag im Verfahrensverzeichnis abgelegt (vgl. [Una16, § 7 Abs. 1 LDSG] und [Bie+18, Annex II]).

Da dem Helpdesk während der Testung eine Sonderrolle zukommt, werden die Testzeiträume auch mit den Mitarbeitern des Helpdesks abgestimmt. Die Leitung des IT-Betriebs fordert, dass über die gesamte Testung hinweg der Testanwender vor Ort ist. Diese Forderung verkürzte den ursprünglich geplanten Gesamtmesszeitraum von mehreren Wochen auf mehrere Tage. Hier entsteht ein Kompromiss zwischen dem Stress, dem die Probanden durch die Messung ausgesetzt sind sowie dem damit einhergehenden möglichen Einfluss auf die Messergebnisse (vgl. Bedingung 7 auf Seite 59), und der Durchführung an sich. Die Testzeiten werden nach Absprache mit dem Gesamtpersonalrat an den vor

7-Zip 9.20		
Adobe Flash Player 10 ActiveX	Adobe Systems Incorporated	10.3.183.5
Adobe Reader X (10.1.0) - Deutsch	Adobe Systems Incorporated	10.1.0
Citrix Online Plug-in	Citrix Systems, Inc.	12.1.44.1
DHTML Editing Component	Microsoft Corporation	6.02.0001
Engineering Client Viewer 7.0	SAP AG	
IrfanView (remove only)	Irfan Skiljan	4.30
Java(TM) 6 Update 30	Oracle	6.0.300
McAfee Agent	McAfee, Inc.	4.8.0.1500
McAfee VirusScan Enterprise	McAfee, Inc.	8.8.04001
Microsoft .NET Framework 4 Client Profile	Microsoft Corporation	4.0.30319
Microsoft .NET Framework 4 Client Profile DEU La...	Microsoft Corporation	4.0.30319
Microsoft Access Runtime 2010	Microsoft Corporation	14.0.4763.1000
Microsoft Office 2010 Primary Interop Assemblies	Microsoft Corporation	14.0.4763.1150
Microsoft Office Standard 2010	Microsoft Corporation	14.0.7015.1000
Microsoft redistributable runtime DLLs VS2005 SP1...	SAP	8.0.50727.4053
Microsoft redistributable runtime DLLs VS2008 SP1...	SAP AG	9.0
Microsoft redistributable runtime DLLs VS2010 SP1...	SAP	10.0.40219.1
Microsoft Visio Viewer 2010	Microsoft Corporation	14.0.4763.1000
Microsoft Visual C++ 2010 x86 Redistributable - ...	Microsoft Corporation	10.0.40219
Microsoft Visual Studio 2010 Tools for Office Runti...	Microsoft Corporation	10.0.50903
Microsoft Visual Studio 2010-Tools für Office-Lauf...	Microsoft Corporation	10.0.50903
NetSupport Notify	NetSupport Ltd	2.01.0004
PDFCreator	pdfforge	1.7.1
Realtek High Definition Audio Driver	Realtek Semiconductor Corp.	6.0.1.6582
roXtra EditClient	Rossmanith GmbH	6.1.0805
SAP Business Explorer	SAP AG	7.30
SAP BusinessObjects Analysis, edition for Microso...	SAP AG	1.4
SAP GUI for Windows 7.30	SAP AG	7.30 Compilation 3
Shared Add-in Extensibility Update for Microsoft	Microsoft	1.0.0
Synaptics Pointing Device Driver	Synaptics Incorporated	17.0.18.8
VLC media player 1.1.11	VideoLAN	1.1.11

Abbildung 18: *Bildschirmaufnahme der in der Standardkonfiguration installierten Software eines Bildschirmarbeitsplatzes, 2016. Zeigt die Spalten „Name“, „Herausgeber“ und „Version“.*

Ort üblichen Bürozeiten ausgerichtet – Montag bis Freitag 9:00 bis 12:00 und 13:00 bis 16:00 Uhr.

Probandenselektion

Aus technischen und ethischen Gründen ist die Messung im Rahmen des Projekts auf die Mitarbeiter der Verwaltung beschränkt. Eine entsprechende Dienstvereinbarung wurde mit dem zuständigen Personalrat getroffen (vgl.

[Bie+18, Annex I] und Abschnitt 4.2.2). Alle zur Messung ausgewählten Artefakte wurden dem Personalrat vor Abschluss vorgestellt. Sie finden jedoch keine explizite Erwähnung in der Dienstvereinbarung.

Um Bedingung 2 auf Seite 56 zu erfüllen, werden ausschließlich Probanden aufgenommen, die einen gültigen Arbeitsvertrag mit dem Betreiber der Infrastruktur haben. Aus dieser Tatsache ergibt sich, dass jeder potenzielle Proband in der Lage ist, der Studienteilnahme einzuwilligen (vgl. [BGB20, § 113]). Der Testanwender erhält eine Liste der Probanden mit den erforderlichen Beschreibungen aus der Personalabteilung.

6.1.2 Pretest

Der Pretest findet vor der Intervention statt und dient der Erfassung der Ausgangslage. Für den Pretest ist der Zeitraum einer Arbeitswoche vorgesehen. In diesem Zeitraum werden acht Einzeltests durchgeführt (vgl. Bedingung 1 auf Seite 48). Die zum Einsatz kommenden Artefakte sind in Tabelle 4 verzeichnet. Die Artefakte wurden nach den in Kapitel 3 erarbeiteten Prinzipien gestaltet. Insbesondere ist Bedingung 11 zu beachten. Eine detaillierte Beschreibung ist Anhang A zu entnehmen.

Tabelle 4: *Artefakte des Pretests.*

Artefakt	Klasse	Beschreibung (Seite)
Selbstlöschende Datei	statisch hostbasiert	119
Updater	dynamisch hostbasiert	120
Targo Bank	gerichtet kommunikationsbasiert	122
Nur Link	gerichtet kommunikationsbasiert	124
IT Ticket	gerichtet kommunikationsbasiert	124
Anti Virus	dynamisch hostbasiert	126
Defacing	ungerichtet kommunikationsbasiert	128
IT Department	gerichtet kommunikationsbasiert	129

6.1.3 Nachbesprechung

Direkt nach Abschluss der Messphase werden die Probanden persönlich per E-Mail kontaktiert (vgl. Bedingung 4 auf Seite 56). Sie werden über den Umfang und das Ziel des Experiments informiert. Ihnen wird eine Ansprechperson benannt, die alle Fragen zum Experiment und dem Forschungsprojekt beantworten kann. Darüber hinaus wird den Probanden angeboten, ihre Daten zu löschen und sie von der zweiten Messung auszuschließen (vgl. Bedingung 3 auf Seite 56).

Diese Nachricht wird mit einer Einladung zu einem Fragebogen, wie in Abschnitt 2.3 vorgeschlagen, begleitet. Der Fragebogen wurde jedoch von einem Partner innerhalb des Forschungsprojekts (vgl. Abschnitt 1.2) erstellt und ausgewertet. Die Ergebnisse sind aus diesem Grund nicht Teil dieser Dissertation. Die Ergebnisse sind bis zum Zeitpunkt des Verfassens dieser Arbeit nicht veröffentlicht.

Darüber hinaus wird allen Probanden eine für sie kostenfreie Teilnahme an einer Intervention angeboten (vgl. Bedingung 9 auf Seite 60). Diese wird im folgenden Abschnitt beschrieben.

6.1.4 Die Intervention

Zur Erhöhung des IT-Sicherheitsbewusstseins der Probanden wird eine Intervention durchgeführt. Diese wird als Präsenzschulung konzipiert, weil sich diese als besonders effektiv zeigen (siehe Abschnitt 2.3). Präsenzschulungen haben darüber hinaus den Vorteil, dass die Teilnahme der Probanden leicht über eine Teilnehmerliste zu dokumentieren ist und jeder der Teilnehmer kontrolliert im gleichen Umfang teilnimmt. Dieser Fakt ergibt sich aus den Dokumentationspflichten im Rahmen von gängigen Zertifizierungsprozessen in der IT-Sicherheit (vgl. Abschnitt 1.1).

Das Gestaltungsziel der Intervention ist eine Schulung der Probanden „in verschiedenen Kompetenzbereichen“ [Mül17]. Die Schulung sieht einen Block aus einem 60-minütigen Vortrag eines Dozenten vor. Der Vortrag behandelt fünf Themenkomplexe: E-Mails, Passwörter, Browser, physische Medien und Smartphones. Insbesondere wird der sichere Umgang mit diesen Technologien

hervorgehoben. Der Vortrag wurde medial mit einem Foliensatz aus 48 Folien unterstützt. Im Rahmen des Vortrags wurde der als „BadUSB“ bekannte Angriff [NL14] demonstriert. Nach dem Vortrag gab es für die Teilnehmer 15 Minuten lang die Gelegenheit, eigene Fragen zu stellen. Das Protokoll dokumentiert eine rege Teilnahme (vgl. [TM17, Abschnitt 3.2.6]). Abschließend wurden die Inhalte, 15 Minuten lang durch den Dozenten, zusammengefasst.

Die Schulung wurde von einem Partner innerhalb des Forschungsprojekts (vgl. Abschnitt 1.2) gestaltet und durchgeführt. Für eine frei zugängliche Veröffentlichung des Schulungsmaterials liegt keine Freigabe vor. Die Ergebnisse sind aus diesem Grund nicht Teil dieser Dissertation.

Begleitend zu der Durchführung der Schulung wird eine E-Learning-Plattform veröffentlicht. Hier wird zu jedem Thema ein Video angeboten, in dem die Schulungsthemen wiederholt werden. Ein Video ist im Schnitt 9:19 Minuten lang ($\sigma = 2{:}01$, $min = 6{:}11$, $max = 10{:}43$). Angeschlossen an das Video ist ein Multiple-Choice-Test aus 10 Fragen zu den Inhalten des Videos. Fehlerhafte Antworten werden bei Abschluss des Tests als solche kommuniziert und der Nutzer hat die Gelegenheit, seine Antworten zu korrigieren. Auf der E-Learning-Plattform war personalisiertes Monitoring der Teilnehmer nicht möglich.

Da die Mitarbeiter des Anwendungspartners über zwei Standorte verteilt sind, wird die Schulung nach der ersten Messphase einmal pro Standort an einem festen Datum angeboten. Die Möglichkeit zur Teilnahme ist durch die Raumgröße und die angebotenen Termine beschränkt.

6.1.5 Posttest

Die zweite Messphase, der Posttest, ist auch über fünf Werktage angelegt. Die zum Einsatz kommenden Artefakte sind in Tabelle 5 (Seite 93) verzeichnet. Diese sollten sich in ihrer Gestalt von den Artefakten des Pretest unterscheiden. Wie im Pretest ist auch hier Bedingung 11 zu beachten. Eine detaillierte Beschreibung ist Anhang A zu entnehmen.

Es wird angenommen, dass die Intervention das IT-Sicherheitsbewusstsein der Probanden erhöht. Da diese sich insbesondere bei Wahrnehmung von Angriffen zeigen soll und diese in ihrer Gestalt variieren (vgl. Abschnitt 3.2), sollte auch der Posttest vom Pretest abweichen. Dies schafft die Angleichung der

Nutzersituation bezüglich der Erwartung des Artefakts und wirkt einem direkten Wiedererkennen entgegen.

Tabelle 5: *Artefakte des Posttests.*

Artefakt	Typ	Beschreibung (Seite)
Versichertenkarte	gerichtet kommunikationsbasiert	130
Java Update	dynamisch hostbasiert	132
Login-Fenster	statisch hostbasiert	127
Word Macro	statisch hostbasiert	133
Bounce	gerichtet kommunikationsbasiert	135

6.2 Durchführung

Der Pretest wurde in einem Zeitraum von 5 Tagen, vom 17. Juli bis zum 21. Juli 2017, durchgeführt. Alle 196 Mitarbeiter aus der Klinikverwaltung wurden nach den Kriterien aus Abschnitt 6.1.1 als Probanden ausgewählt. In der Gruppe der Probanden existiert ein Verhältnis von 4,6 Frauen zu einem Mann. Abbildung 19 zeigt die relative Verteilung der Probanden über die Abteilungen der Klinikverwaltung. Die Daten über die Probanden wurden nur anonymisiert zur Verfügung gestellt. Eine Auswertung nach diesen Daten ist deshalb nicht möglich.

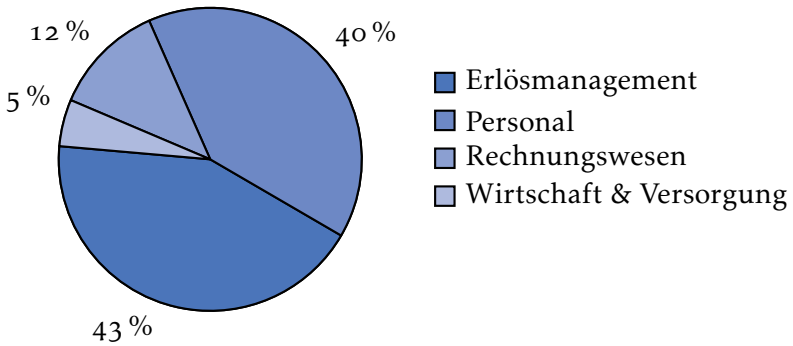


Abbildung 19: *Verteilung der Probanden nach Abteilung.*

Nach der Probandenselektion liegen keine weiteren Informationen über die Probanden vor. Aufgrund der Automatisierung der Messung nehmen genau die Probanden tatsächlich an der Messung teil, die sich im Testzeitraum an einem Bildschirmarbeitsplatz anmelden, auf dem die Clientsoftware installiert wurde. Ein Test der ersten Messphase wurde im Durchschnitt von 51,25 Probanden durchlaufen ($\sigma = 4{,}27$, min = 44, max = 57). Ein Proband durchläuft während des Posttests im Durchschnitt 5,47 Tests ($\sigma = 2{,}31$, min = 1, max = 8).

Im Anschluss an die erste Messphase wurde eine E-Mail an alle 196 Probanden der ersten Phase zur Nachbesprechung versandt (vgl. Abschnitt 6.1.3). Insgesamt gingen 102 Anfragen nach der Schulung ein. 52 Probanden meldeten sich zu den angebotenen Terminen (21. und 22. September 2017) für die Schulung an. Diese Probanden nahmen auch an der Schulung teil. Von den weiteren 50 Anfragen konnten jedoch nur 38 der Gruppe der Probanden zugeordnet werden. Aufgrund der hohen Nachfrage wurde ein weiterer Schulungstermin nach der zweiten Messphase angeboten.

Vier Probanden wünschten eine Löschung ihrer Daten und die Exklusion von weiteren Versuchen. Die Daten der Probanden wurden aus dem Datensatz entfernt und finden keine Berücksichtigung in den vorliegenden Betrachtungen. Sie nehmen nicht an der zweiten Messphase teil.

Die zweite Messphase wurde zwischen dem 2. und 7. Mai 2018 umgesetzt. Damit liegt sie etwa zehn Monate nach der ersten Messphase und etwa sieben Monate nach der Intervention. Alle 163 Nutzer von Bildschirmarbeitsplätzen in der Klinikverwaltung wurden zur Teilnahme an dieser zweiten Phase ausgewählt. Die relative Verteilung der Probanden auf die Abteilungen der Klinikverwaltung folgt wieder der in Abbildung 19 gezeigten. Das Verhältnis von Frauen zu Männern ist jedoch mit 5 zu 1 ein wenig verschoben. Die vier Probanden, die keine Teilnahme an weiteren Messphasen wünschen, sind ausgeschlossen.

104 Probanden wurden für beide Messphasen selektiert. 60 Probanden nahmen erstmals im Posttest an einer Messung teil. Entsprechend nahmen von den 196 selektierten Probanden 92 nur am Pretest teil. Da dieselbe Selektionsstrategie in beiden Phasen Anwendung findet, ist die Diskrepanz in der Anzahl der Probanden nur durch die Fluktuation in der Belegschaft zu erklären.

Innerhalb der zweiten Messphase wird ein Einzeltest im Durchschnitt von 80 Probanden durchlaufen ($\sigma = 44{,}90$, min = 36, max = 132). Dabei durchlief ein Proband im Durchschnitt 2,47 Tests ($\sigma = 0{,}89$, min = 1, max = 5).

Auch nach der zweiten Messphase wurde eine E-Mail an die Probanden der Messphase verschickt. Kein Proband wünschte von seinem Recht von der Studienteilnahme zurückzutreten, Gebrauch zu machen. Da nicht alle Wünsche nach der Teilnahme an der Schulung nach dem Pretest berücksichtigt werden konnten, wurde die Schulung erneut angeboten (vgl. Bedingung 9). Die zweite Durchführung wurde nicht protokolliert.

6.3 Ergebnisse

In diesem Kapitel werden die erfassten Nutzerreaktionen ausgewertet. Zunächst werden die erfassten Reaktionsdaten vorgestellt. Auf Grundlage dieser Daten wird die Validitätsprüfung über eine Interventionseffektmessung angestellt. Diese wird in Abschnitt 6.3.2 beschrieben. Da es sich bei der Selektion der Probanden für die Intervention um eine Form von Selbstselektion handelt, ist denkbar, dass die Motivation und das Interesse der Probanden allein einen großen Einfluss auf den gemessenen Interventionseffekt haben. Dies wird in Abschnitt 6.3.3 untersucht.

Da in diesem Experiment erstmals die gleichen Probanden mit mehreren Artefakten konfrontiert werden und Belege für die Reliabilität von verhaltensbasierten Feldexperimenten bisher nicht vorliegen (vgl. Abschnitt 2.2.4), wird die Reliabilität dieser Methode untersucht. Die hohe Varianz der erfassten Reaktionsdaten macht eine Abschätzung der Reliabilität der Messungen mit Artefakten unabdingbar. Diese Untersuchung wird in Abschnitt 6.3.4 dokumentiert. Abschnitt 6.3.5 illustriert die Implikationen der erfassten Daten in Bezug auf das IT-Sicherheitsbewusstsein von Gruppen, wie in Abschnitt 3.4.2 beschrieben.

6.3.1 Erfasste Reaktionen

Die im Pretest erfassten Reaktionen der Probanden sind in Abbildung 20 auf der nächsten Seite dargestellt. Auffällig ist, dass die Interaktionsraten bei den gerichteten kommunikationsbasierten Artefakten deutlich geringer sind als

bei den Artefakten anderer Klassen. Darüber hinaus wird deutlich, dass die erfassten Reaktionen stark schwanken. Diese Beobachtung unterstützt die bereits in Bedingung 1 (Seite 48) formulierte Forderung nach mehreren Einzeltests pro Proband. Darüber hinaus machen diese Zahlen deutlich, dass ein direkter Vergleich der Artefakte aufgrund ihrer technischen Gestaltung nicht möglich ist. Eine Untersuchung der Reliabilität verhaltensbasierter Experimente wird weiter motiviert.

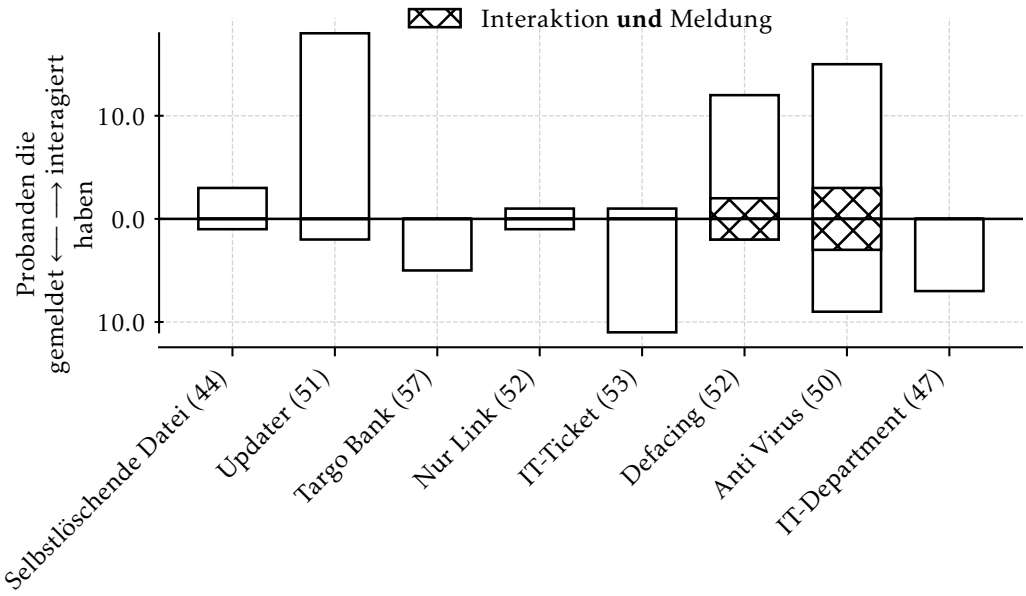


Abbildung 20: *Im Pretest aufgezeichnete Probandenreaktionen nach ihrem Typ pro Artefakt. Nach oben der Anteil der Probanden, der mit dem Artefakt interagiert hat, nach unten der Anteil der Probanden, der das Artefakt gemeldet hat. Probanden, die beide Handlungsoptionen wahrgenommen haben, sind schraffiert dargestellt. In Klammern ist die Anzahl der Probanden notiert, an die das Artefakt ausgerollt wurde (vgl. Anhang A).*

In Abbildung 21 sind die erfassten Reaktionen der Probanden des Posttests auf die Artefakte aufgetragen. Besonders auffällig ist die hohe Anzahl an Reaktionen für die Artefakte `java-update` und `login-window`. Dies ist einer besonderen Kombination aus Umständen geschuldet. Zum einen handelt es sich dabei um dynamische hostbasierte Artefakte. Als solche werden diese nicht-konditional in das Wahrnehmungsfeld der Probanden injiziert. Damit haben sie ein hohes Potenzial eine Interaktion zu provozieren.

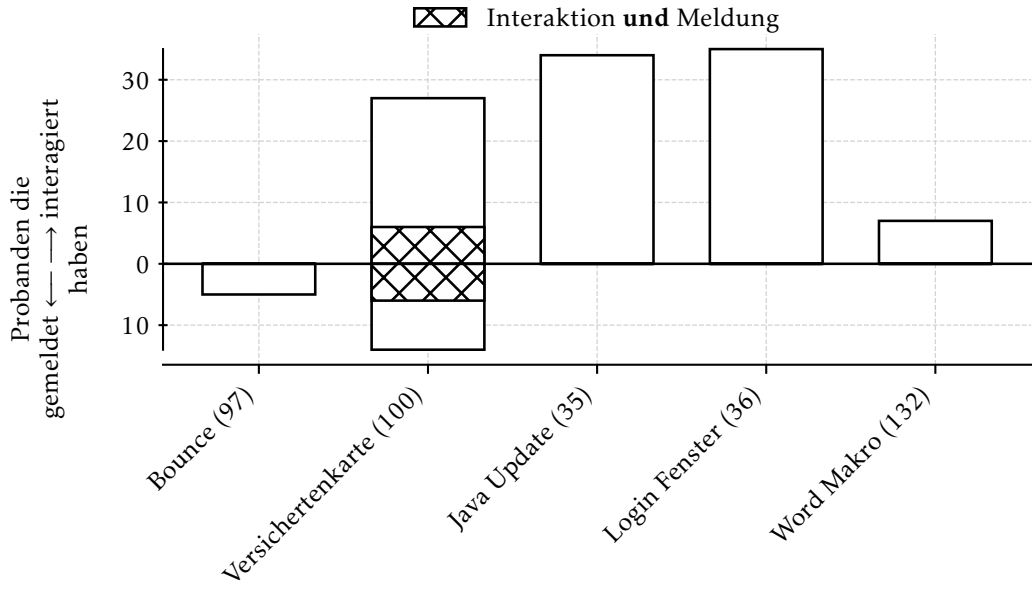


Abbildung 21: *Im Posttest aufgezeichnete Probandenreaktionen nach ihrem Typ pro Artefakt. Nach oben der Anteil der Probanden, der mit dem Artefakt interagiert hat, nach unten der Anteil der Probanden, der das Artefakt gemeldet hat. Probanden die beide Handlungsoptionen wahrgenommen haben, sind schraffiert dargestellt. In Klammern ist die Anzahl der Probanden notiert, an die das Artefakt ausgerollt wurde (vgl. Anhang A).*

Diese Tatsache liefert jedoch keine Erklärung dafür, warum in der ersten Messphase die dynamischen hostbasierten Artefakte keinen derart extremen Ausschlag in Artefaktinteraktion bei den Probanden erzeugen. Dies liegt an der Weiterentwicklung des Messwerkzeugs. Während das Werkzeug im Pretest noch vorausgesetzt hat, dass der Proband zu Beginn des Testzeitraums angemeldet ist, genügt es im Posttest, dass der Proband zu einem beliebigen Zeitpunkt in dem Testzeitraum angemeldet ist. Dadurch ergibt sich mehr Gelegenheit dafür, dass das Reaktionspotenzial des Artefakts (vgl. Abschnitt 3.2.5) zum Tragen kommt.

Tabelle 6 auf der nächsten Seite zeigt eine Übersicht über die Teilnahme der Probanden an den verschiedenen Komponenten des Experiments. Eine Teilnahme an einem Teil des Experiments ist mit „✓“ gekennzeichnet. Ein „✗“ steht für das Gegenteil. In der Tabelle werden die Teilnahme an den Messphasen, die Teilnahme an der Intervention, aber auch die bekundete Motivation der

Probanden codiert. Ein Proband hat an einer Messphase teilgenommen, wenn mindestens ein Artefakt dieser Phase an ihn ausgeliefert wurde. Ein Proband ist motiviert, wenn er seinen Wunsch zur Teilnahme an der Intervention bekundet hat. Demnach kann ein Proband nicht an der Intervention teilgenommen haben, aber unmotiviert sein. Diese Kombinationen sind nicht in der Tabelle verzeichnet.

Entgegengesetzt zur Darstellung in einer vorhergehenden Arbeit [Syk+20] verzeichnet diese Darstellung auch den Anteil der Probanden, die von Beginn des Experiments an diesem teilnehmen. Diese sind anhand ihres Pseudonyms zu erkennen. Probanden mit Pseudonymen, deren *id* kleiner oder gleich 196 ist (siehe dazu Abbildung 16 auf Seite 83), wurden bereits für die erste Messphase selektiert. Hier fällt auf, dass zur zweiten Phase 60 Probanden in das Experiment eingetreten sind. 59 davon haben mindestens einen Test der zweiten Phase durchlaufen. Ein Proband aus dieser Gruppe hat keinen Test durchlaufen.

Tabelle 6: *Größe der sich durch die Durchführung der Studie ergebenden Gruppen von Probanden und derer Kombinationen.* ✓ *codiert die Teilnahme der Probanden an einem Experimentteil,* ✗ *codiert das Ausbleiben der Teilnahme. Die Experimentteile sind: der Pretest* P1*, die Motivation* M*, die Intervention* I*, und der Posttest* P2*. Ein Proband* P *hat an einer Messphase teilgenommen, wenn mindestens ein Artefakt dieser Phase an ihn ausgeliefert wurde. Die Spalte* Gruppe *codiert in welche Vergleichsgruppe Probanden aufgrund ihrer Teilnahme fallen.*

P1	M	I	P2	#P	#P(id≤196)	Gruppe
✓	✓	✓	✓	3	3	A, X
✓	✓	✓	✗	22	22	X
✓	✓	✗	✓	7	7	B, Y, K
✓	✓	✗	✗	4	4	Y
✓	✗	✗	✓	26	26	B, Y, J
✓	✗	✗	✗	13	13	Y
✗	✓	✓	✓	18	18	A
✗	✓	✓	✗	9	9	
✗	✓	✗	✓	14	14	B, K
✗	✓	✗	✗	13	13	
✗	✗	✗	✓	94	35	B, J
✗	✗	✗	✗	33	32	

Nur 3 Probanden haben an allen Teilen des Experiments teilgenommen. Das unterstreicht die Unabhängigkeit der Stichproben der Messphasen. Je nach Untersuchung in den folgenden Abschnitten werden die Probanden, die zum

Vergleich herangezogen werden, gruppiert. Die letzte Spalte in Tabelle 6 verzeichnet die Gruppenzugehörigkeit der Probandengruppe in den folgenden Betrachtungen. An den entsprechenden Stellen werden die Teilnahmen der Gruppen referenziert. Die Darstellung „✓ – – ✗“ bedeutet: „Alle Probanden, die an der ersten Messphase teilgenommen haben, ungeachtet der Motivation oder Teilnahme an der Intervention, ohne Teilnahme an der zweiten Testphase“.

6.3.2 Demonstration der Validität

Um die Validität der Methode zur IT-Sicherheitsbewusstseinsmessung auf Basis von Artefakten zu demonstrieren, wird die Effektstärke einer Intervention gemessen, von der aktuell angenommen wird, dass sie einen großen Effekt hat (vgl. dazu Abschnitt 1.1 und Abschnitt 2.3). Es wird das IT-Sicherheitsbewusstsein der Probanden, die vor der zweiten Messung an der Intervention teilgenommen haben, mit denen, die nicht an dieser teilgenommen haben, verglichen. Damit ergibt sich für diesen Vergleich die Teilnahme an der Schulung als unabhängige Variable und das in der zweiten Phase demonstrierte individuelle IT-Sicherheitsbewusstsein als abhängige Variable. Lässt sich eine Steigerung in dem IT-Sicherheitsbewusstsein der Probanden zeigen, die an der Intervention teilgenommen haben, ist anzunehmen, dass die Messung valide ist.

Abbildung 22 auf der nächsten Seite zeigt für das demonstrierte individuelle IT-Sicherheitsbewusstsein ein Histogramm und zur Illustration die über eine Kerndichteschätzung [Sco15] approximierte Verteilung. Verglichen wird dieser Wert zwischen den Probanden, die an der Intervention teilgenommen haben (Gruppe A in Tabelle 6: – – ✓ ✓, $N = 21$), und denen, die nicht an dieser teilgenommen haben (Gruppe B in Tabelle 6: – – ✗ ✓, $N = 82$). Bei dem angestellten Vergleich werden Probanden unabhängig ihrer Teilnahme am Pretest und der dokumentierten Motivation zusammengefasst. Die Aufklärung, der Fragebogen und die Einladung zur Schulung wurden an alle diese Probanden versendet. Ihr Pseudonym hat eine $id \leq 196$. Die Artefakte werden zu keinem Zeitpunkt gegenüber den Probanden thematisiert. Damit ergibt sich als einziger Unterschied zwischen den Probanden die Möglichkeit, ein Artefakt wahrgenommen zu haben. Aus diesem Grund ist der Einfluss der Teilnahme am Pretest selbst als gering einzuschätzen.

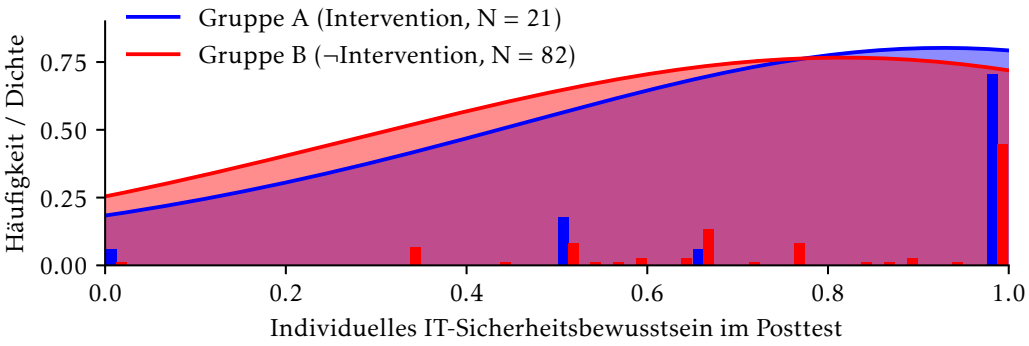


Abbildung 22: *Häufigkeit und geschätzte Dichte des durch Probanden im Posttest demonstrierten individuellen IT-Sicherheitsbewusstseins. Gruppe A (blau) hat zuvor eine Intervention durchlaufen. Gruppe B (rot) hat keine Intervention durchlaufen.*

Zu sehen ist, dass die Probanden, die die Intervention durchlaufen haben, ein leicht höheres individuelles IT-Sicherheitsbewusstsein (0,86) demonstrieren als die Probanden, die nicht an der Intervention teilgenommen haben (0,77). Damit zeigt sich eine Erhöhung von 11,69 %.

Zur Prüfung, ob eine Steigerung tatsächlich statistisch nachvollziehbar ist, kommt ein zweiseitiger Mann-Whitney-U-Test [MW47] zum Einsatz. Er erlaubt eine Aussage darüber, ob ein in zwei Stichproben beobachteter Unterschied aus der gleichen Grundgesamtheit stammen kann und damit die Abweichung *zufällig* aufgetreten ist. Die Testgröße $p = 0{,}12$ lässt das Verwerfen der Nullhypothese, also dass die Interventionsgruppe kein zufällig höheres IT-Sicherheitsbewusstsein zeigt, bei einem klassisch gewählten Signifikanzniveau von 0,05 nicht zu. Damit lässt sich eine Steigerung des IT-Sicherheitsbewusstseins durch Effektstärke nicht abschließend demonstrieren.

Wird der Vergleich zwischen Teilnehmern des Pretests angestellt, so zeigt sich jedoch kein Unterschied zwischen den Gruppen. Für diesen Vergleich wird das individuelle IT-Sicherheitsbewusstsein der Probanden, die eine Intervention erhalten werden, mit denen verglichen, die nicht an der Intervention teilnehmen werden. Dieser Vergleich wird ohne Achtung der Motivation angestellt und folgt damit dem gleichen Prinzip des zuvor angestellten Vergleich.

Die Gruppe der Probanden, die eine Intervention erhalten wird (Gruppe X in Tabelle 6: ✓ – ✓ –, $N = 25$), demonstriert in der ersten Messphase ein leicht geringeres individuelles IT-Sicherheitsbewusstsein, im Durchschnitt 0,89, als die Probanden, die nicht an der Intervention teilnehmen werden (Gruppe Y in Tabelle 6: ✓ – ✗ –, $N = 50$). Letztere demonstrieren auf den Artefakten des Pretests ein durchschnittliches individuelles IT-Sicherheitsbewusstsein von 0,90. Der zweiseitige Mann-Whitney-U-Test zeigt einen p-Wert von 0,95. Damit sollte die Nullhypothese nicht verworfen werden. Abbildung 23 zeigt die Verteilung der Werte analog zu Abbildung 22.

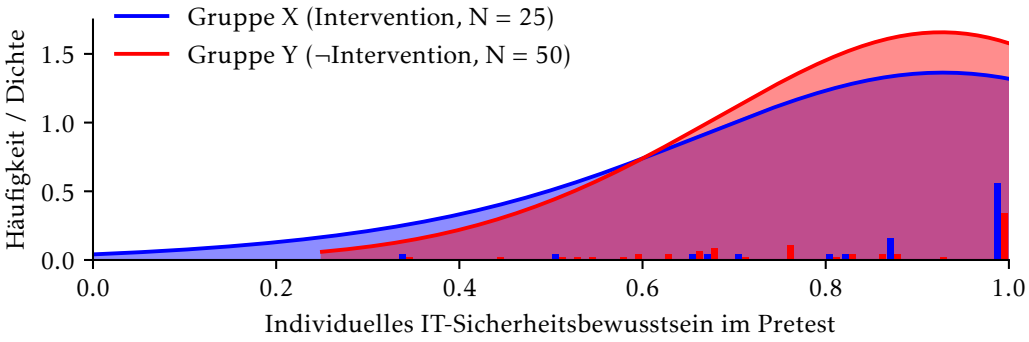


Abbildung 23: *Häufigkeit und geschätzte Dichte des durch Probanden im Pretest demonstrierten individuellen IT-Sicherheitsbewusstseins. Gruppe X (blau) wird darauf folgend eine Intervention durchlaufen. Gruppe Y (rot) wird die Intervention nicht durchlaufen.*

6.3.3 Einfluss der Motivation

Da die Selektion der Probanden für die Intervention durch eine Einladung erfolgt ist (vgl. Abschnitt 6.1.4), ist möglich, dass eine Verzerrung durch Selbstselektion aufgetreten ist. Es ist denkbar, dass Probanden, die zur Teilnahme einer Schulung zur Erhöhung des IT-Sicherheitsbewusstseins motiviert sind, generell ein höheres IT-Sicherheitsbewusstsein demonstrieren. Es war nicht möglich, alle teilnahmebereiten Probanden die Intervention durchlaufen zu lassen. Die Termine zur Durchführung der Intervention sowie die räumliche Kapazität sind limitiert. Dadurch ergibt sich ein pseudo-zufälliger Einfluss auf die Selektion. Der Vergleich des demonstrierten IT-Sicherheitsbewusstseins des Posttests der motivierten Probanden, die nicht an der Schulung teilgenommen haben, mit den unmotivierten Probanden gibt Aufschluss über den Effekt

der Motivation. Damit ergibt sich für diesen Vergleich die dokumentierte Motivation zur Teilnahme an der Schulung als unabhängige Variable und das in der zweiten Phase demonstrierte individuelle IT-Sicherheitsbewusstsein als abhängige Variable. Wie in den Vergleichen zuvor wird die Teilnahme an dem Pretest nicht berücksichtigt.

Es ergibt sich eine Gruppe aus Probanden, die zur Teilnahme bereit ist, denen die Teilnahme jedoch aufgrund unkontrollierter Bedingungen versagt ist, Gruppe K ($N = 21$) in Tabelle 6: – ✓ ✗ ✓. Der Vergleich des im Posttest demonstrierten individuellen IT-Sicherheitsbewusstseins dieser Probanden mit dem der Probanden, die sich nicht zur Teilnahme bereit erklärt haben, Gruppe J ($N = 61$) in Tabelle 6: – ✗ ✗ ✓, gibt Aufschluss über Selektionsverzerrungen.

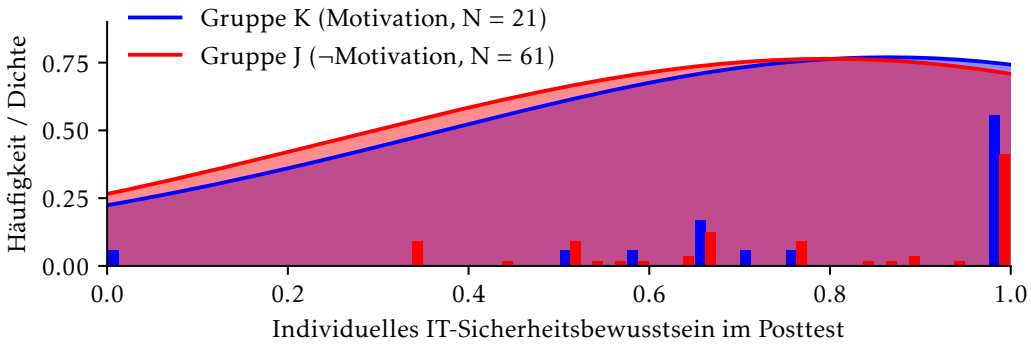


Abbildung 24: *Häufigkeit und geschätzte Dichte des durch Probanden im Posttest demonstrierten individuellen IT-Sicherheitsbewusstseins. Gruppe K (blau) hat zuvor den Wunsch nach der Teilnahme an einer Intervention zum Ausdruck gebracht, Gruppe J (rot) hat das nicht. Keine Gruppe hat die Intervention durchlaufen.*

Gruppe K zeigt einen mittleren Wert für das individuelle IT-Sicherheitsbewusstsein von 0,81 und Gruppe J von 0,76. Ein zweiseitiger Mann-Whitney-U-Test liefert $p = 0,42$. Damit kann ein signifikanter Effekt der Motivation auf das IT-Sicherheitsbewusstsein nicht nachgewiesen werden.

6.3.4 Reliabilität

Die Durchführung hat gezeigt, dass die Raten, mit denen Probanden auf unterschiedliche Artefakte reagieren, stark variieren (vgl. Abbildungen 20 und 21). Bisherige akademische Dokumentationen von Feldexperimenten in

diesem Bereich testen lediglich einzelne Artefakte oder zeigen ebenso eine hohe Variabilität der „*click rate*“ (vgl. Abschnitt 2.3).

Aus diesen Gegebenheiten erwächst die Frage, ob Probanden in artefaktbasierten Tests überhaupt eine Verhaltenstendenz zeigen. Aus diesen Gründen sind zwei Fragen zu beantworten:

1. Wie konsistent sind Messungen von IT-Sicherheitsbewusstsein anhand von Phishing-E-Mails?
2. Wie konsistent sind die Messungen von IT-Sicherheitsbewusstsein anhand von Nicht-Phishing-E-Mail-Artefakten?

Um diese Fragen zu beantworten, wird die *interne Konsistenz* [DB16] der Messung in verschiedenen Klassen berechnet. Die interne Konsistenz gibt an, wie stark das Verhalten der Probanden bei unterschiedlichen Artefakten miteinander korreliert. Mathematisch modelliert wird diese Korrelation mit dem Pearson-Korrelationskoeffizient ρ. Damit ergibt sich für zwei Artefakte a, a' gleicher Klasse und allen Probanden s, die an beiden Tests teilgenommen haben, die in Gleichung (6.1) abgebildete Berechnungsvorschrift. Diese basiert auf dem Maß für individuelles IT-Sicherheitsbewusstsein, wie in Abschnitt 3.4.1 beschrieben.

$$\rho(P_{s,a}(\{i \wedge \neg r\}),\ P_{s,a'}(\{i \wedge \neg r\})) \tag{6.1}$$

Abbildung 25 auf der nächsten Seite zeigt einen Box-Plot der Korrelationskoeffizienten. Die Klassen sind „Phishing“ aus 5 Artefakten und „¬Phishing“ aus 8 Artefakten (vgl. Tabellen 4 und 5). Diese Einteilung ist zum einen der Beliebtheit von Phishing-E-Mails in den verwandten Arbeiten geschuldet, aber auch der Tatsache, dass den Probanden das Angriffsszenario, auch in der durchgeführten Intervention, mit besonderem Fokus vermittelt wird. Verglichen werden Artefakte innerhalb der jeweiligen Klasse, unter den Klassen und gemischt.

Tabelle 7 auf der nächsten Seite zeigt die entsprechenden Maßzahlen in der Übersicht. Um diese Werte zu kontextualisieren, wird das Cronbachsche α berechnet [DB16, S. 468]. Insbesondere auffällig präsentiert sich, dass fast alle ermittelten Werte für ρ positiv sind. Das deutet darauf hin, dass es sich um die Vermessung des gleichen Konzepts handelt. Auch liegt ρ selten in einem

besonders hohen Wertebereich ($\rho \geq 0,7$), was darauf hindeutet, dass keine starke statistische Abhängigkeit zwischen den Werten existiert.

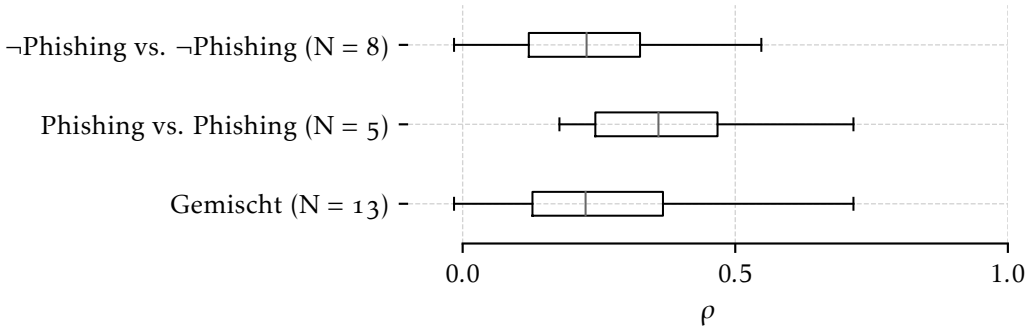


Abbildung 25: *Interne Konsistenz von der Messung mit Artefakten nach Artefaktklasse. Pearson-Korrelationskoeffizient ρ eines paarweisen Vergleichs für die Wahrscheinlichkeit mit einem Artefakt zu interagieren und dieses nicht zu melden $\{i \wedge \neg r\}$. Gruppiert zum Vergleich von Klassen.*

Werte von $\alpha \geq 0,66$ werden mit „akzeptabel“, „befriedigend“, „genügend“ und „angemessen“ konnotiert [Tab18]. Darüber hinaus ist zu bemerken, dass das Cronbachsche α höhere Werte für höhere N ergibt, was den höheren Wert für „Gemischt“ erklärt. In Bezug auf die eingehende Fragestellung ist zu vermerken, dass die Messung mit Phishing-E-Mails konsistenter ist als die Messung mit anderen Artefakten. Es ist damit anzunehmen, dass die artefaktbasierte Messung von IT-Sicherheitsbewusstsein intern konsistent und damit reliabel ist.

Tabelle 7: *Maßzahlen der internen Konsistenz nach Gruppen von Artefakten. Aufgelistet sind pro Gruppe: die Anzahl der Artefakte N, der mittlere Pearson-Korrelationskoeffizient $\bar{\rho}$, die Standardabweichung σ, das Minimum und das Maximum sowie das sich ergebende Cronbachsche α.*

Klasse	N	$\bar{\rho}$	σ	min	max	α
Phishing	5	0,38	0,17	0,18	0,72	0,66
¬Phishing	8	0,24	0,17	-0,02	0,55	0,66
Gemischt	13	0,23	0,17	-0,02	0,72	0,77

6.3.5 Organisatorische Einbettung

Während sich das individuelle IT-Sicherheitsbewusstsein als Metrik zur Validierung der Messung und zur Bildung von Referenzen eignet, stellt dieses Maß, wie bereits in Abschnitt 3.4.2 dargestellt, eine Einbettung in den organisatorischen Kontext nicht optimal dar.

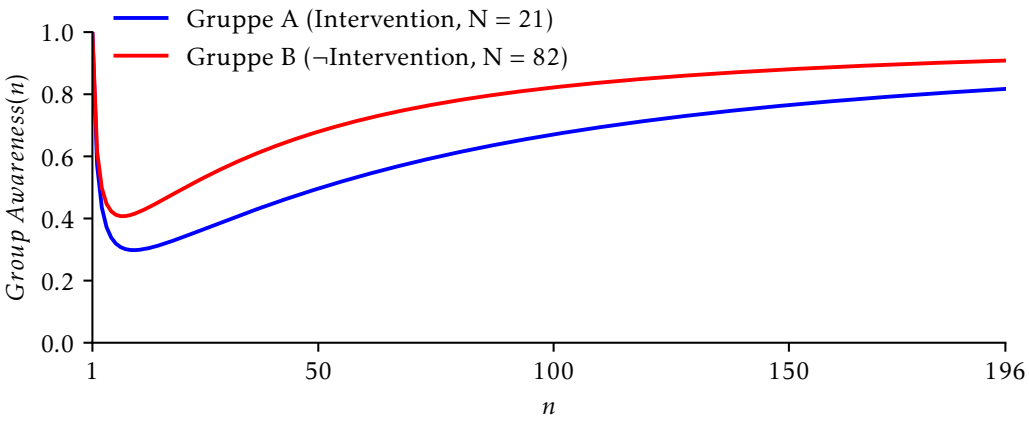


Abbildung 26: *Durch Probanden im Posttest demonstriertes gruppenbasiertes IT-Sicherheitsbewusstsein nach Gleichung* (3.3)*. Gruppe A (blau) hat eine Intervention durchlaufen, Gruppe B (rot) nicht.*

In Abbildung 26 ist das gruppenskalierte IT-Sicherheitsbewusstsein für die Gruppen A (in Tabelle 6: – – ✓ ✓, $N = 21$) und B (in Tabelle 6: – – ✗ ✓, $N = 82$) nach Gleichung (3.3) dargestellt. Dieser Wert ist dabei für alle möglichen Angriffsgrößen n von 1 bis 196, der maximalen Gruppengröße, aufgetragen. Dabei zeigt sich, dass das gruppenbasierte IT-Sicherheitsbewusstsein der Gruppe, die keine Intervention erhalten hat, das der Interventionsgruppe übersteigt, obwohl die Probanden dieser Gruppe ein geringeres individuelles IT-Sicherheitsbewusstsein demonstrieren.

Diese Beobachtung basiert auf dem Einfluss der Wahrscheinlichkeit, mit der ein Proband ein Artefakt meldet, auf den Gesamtwert. Ein erfolgreicher Angriff auf eine Organisation setzt voraus, dass mindestens ein Nutzer mit dem Artefakt interagiert und kein Nutzer, der dieses Artefakt beobachten kann, dieses meldet. Tabelle 8 auf der nächsten Seite zeigt für die jeweiligen Gruppen die Komponenten des in Gleichung (3.3) verwendeten Terms.

Tabelle 8: *Interventionseffekte auf die Komponenten des Maßes für IT-Sicherheitsbewusstsein von Gruppen. Dargestellt sind die Werte für die Interventionsgruppen, Gruppe A bei den Werten P_2 und X bei P_1 sowie den Nicht-Interventionsgruppen, Gruppe B bei P_2 und Y bei P_1.*

Gruppe		$P_2(\{i \wedge \neg r\})$	$P_2(\{\neg r\})$	$P_1(\{\neg r\})$	N
Intervention	A	0,143	0,024	–	21
	X	–	–	0,129	25
¬Intervention	B	0,230	0,043	–	82
	Y	–	–	0,134	24

Die Wahrscheinlichkeit, ein Artefakt zu melden, ist in der Interventionsgruppe, Gruppe A ($N = 21$), geringer (0,024) als in Gruppe B ($N = 82$), die keine Intervention in Anspruch genommen hat (0,043). Diese Differenz zeigt sich nicht zwischen den Gruppen X (in Tabelle 6: ✓ – ✓ –, $N = 25$) und Y (in Tabelle 6: ✓ – ✗ –, $N = 24$).

6.4 Fazit

Im Rahmen dieses Kapitels wurden die Planung und Durchführung des Experiments zur Erprobung der artefaktbasierten Messung von IT-Sicherheitsbewusstsein beschrieben sowie die erfassten Daten analysiert. Damit werden die Antworten auf die Forschungsfragen 2 bis 4 evaluiert.

Die Durchführung berücksichtigt die in Kapitel 4 erarbeiteten Rahmenbedingungen. Tabelle 9 bietet eine Übersicht.

Die im Rahmen dieses Kapitels angestellten Betrachtungen exkludieren die Reaktionsdaten von 59 Probanden, die erst zum Posttest in das Experiment eingetreten sind. Diese erhielten keine Einladung zur Teilnahme an der Intervention und hatten keine Chance zur Teilnahme. Diese Probanden demonstrieren, entgegen der Erwartung [DCF07; BG15], ein überdurchschnittlich hohes IT-Sicherheitsbewusstsein [Syk+20]. Dabei ist jedoch unklar, ob diese Probanden zwischen den Phasen in die Organisation eingetreten sind oder nur ihre Position in dieser verändert haben.

Tabelle 9: *Übersicht der im Rahmen der Durchführung des Experiments berücksichtigten Bedingungen.*

Bedingung	Zusammenfassung
Bedingung 1: Ein Proband muss mehrere Tests mit unterschiedlichen Artefakten durchlaufen.	Sowohl im Pretest (vgl. S. 90) als auch im Posttest (vgl. S. 92) kamen mehrere, unterschiedliche Artefakte zum Einsatz.
Bedingung 2: Kann ein informiertes Einverständnis, beispielsweise durch eine verringerte Geschäftsfähigkeit (vgl. [BGB20, § 104–107]), nicht erteilt werden, sind diese potenziellen Probanden von der Teilnahme auszuschließen.	Es wurden ausschließlich Mitarbeiter der Verwaltung der Klinik selektiert, die einen persönlichen Nutzeraccount haben (vgl. S. 90).
Bedingung 3: Die Erteilung des Einverständnisses muss ohne negative Auswirkungen für den potenziellen Probanden ablehnbar sein.	Den Probanden wurde nach der Information das Recht auf Löschung ihrer Daten eingeräumt (vgl. S. 91).
Bedingung 4: Die Probanden müssen frühestmöglich über das Experiment, ihre Teilnahme, ihre Rechte und wie sie diese in Anspruch nehmen können informiert werden.	Die Probanden wurden nach jeder Messphase persönlich adressiert und umfänglich informiert (vgl. S. 91).
Bedingung 5: Durch den Einsatz des Werkzeugs im Rahmen des Experiments darf kein Nichtproband beeinträchtigt werden, der nicht vorher explizit dem Vorhaben zugestimmt hat.	Bildschirmarbeitsplätze, die in Kontakt zu besonders vulnerablen Personengruppen stehen, wurden nicht in das Experiment aufgenommen (vgl. S. 88).

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 9.

Bedingung	Zusammenfassung
Bedingung 6: Die Wahrscheinlichkeit für eine Beeinträchtigung der IT-Infrastruktur durch den Einsatz des Werkzeugs ist zu minimieren.	Das Werkzeug wurde anhand der in Abschnitt 5.2.1 ausgesprochenen Empfehlungen in Betrieb genommen (vgl. S. 88).
Bedingung 7: Der Zeitraum zur Datenerfassung ist so weit zu strecken, dass ein Proband zu keinem Zeitpunkt einem übermäßig hohen Maß an Stress ausgesetzt ist.	Der Zeitraum der Datenerfassung wurde durch die organisatorische Forderung nach der Anwesenheit des Testanwenders beschränkt. Er wurde jedoch im Rahmen der Umstände maximal gewählt (vgl. S. 89).
Bedingung 9: Jeder Proband soll die gleiche Chance erhalten, an der Schulung teilzunehmen.	Die Probanden wurden nach dem Pretest (vgl. S. 91) und nach dem Posttest zu der Schulung eingeladen (vgl. S. 95).
Bedingung 11: Artefakte dürfen nicht so gestaltet sein, dass „Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung“ [GDP06, Art. 9 Abs. 1] verarbeitet werden.	Die in den Messphasen zum Einsatz kommenden Artefakte wurden entsprechend dieser Bedingung gestaltet (vgl. S. 90 und S. 92 sowie Anhang A).

Die im Rahmen dieses Kapitels beschriebenen Untersuchungsergebnisse konnten einen Interventionseffekt durch die Schulung demonstrieren. Deren Effektstärke

ist jedoch gering. Entsprechend zeigt ein Test auf die Signifikanz eine hohe Testgröße ($p = 0{,}12$), verglichen mit konventionellen Signifikanzniveaus [Lab68]. Dieser Effekt lässt sich jedoch nicht im Pretest feststellen. Die Auswertung des Fragebogens (vgl. dazu Abschnitt 2.3 und Abschnitt 6.1.3) liegt zu diesem Zeitpunkt nicht vor und kann deshalb keinen weiteren Aufschluss bringen. Zwar kann die Steigerung des individuellen IT-Sicherheitsbewusstseins nicht ohne Zweifel dargestellt werden, jedoch deuten die Indizien auf ein Vorhandensein des Effektes hin und motivieren damit weitere Untersuchungen.

Darüber hinaus konnte gezeigt werden, dass die artefaktbasierte Messung von IT-Sicherheitsbewusstsein reliabel ist, jedoch keine besonders hohe Kovarianz zwischen den Reaktionen der Probanden auf ein Artefakt besteht. Auch sind die erfassten Reaktionen stark vom Artefakt abhängig. Die ist insbesondere in Abbildung 20 auf Seite 96 und Abbildung 21 auf Seite 97 illustriert. Dies muss zu der Empfehlung führen, dass verhaltenserfassende Experimente stets mehrere Artefakte einsetzen sollten. Dies muss insbesondere auch für Phishing-Experimente gelten.

Die Untersuchung des IT-Sicherheitsbewusstseins von Gruppen offenbart einen möglichen durch die Intervention entstandenen Missstand. Eine Schulung der Mitarbeiter einer Organisation könnte der IT-Sicherheit der Organisation tatsächlich abträglich sein. Die Existenz dieses Effekts kann jedoch nicht abschließend bewiesen werden und liefert an dieser Stelle lediglich Anreiz für weitere Forschung. Die möglichen Implikationen müssen jedoch zumindest als besorgniserregend klassifiziert werden.

7 Zusammenfassung, Diskussion & Ausblick

Die Leitfrage der vorliegenden Arbeit ist die Frage nach Verbesserungsmöglichkeiten von IT-Sicherheitsbewusstseinsmessungen. Als Antwort auf diese Frage wird in dieser Arbeit ein Konzept zur ethischen und rechtskonformen, nutzerverhaltensbasierten Messung von IT-Sicherheitsbewusstsein hergeleitet, implementiert und erprobt.

Um dies zu erreichen, wurden aus der Leitfrage vier Forschungsfragen abgeleitet. Die erste Forschungsfrage fokussiert bestehende Methoden zur IT-Sicherheitsbewusstseinsmessung und analysiert deren Stärken und Schwächen. Dieser Analyse geht die Definition des zu vermessenden Gegenstandes voraus. Es konnte dargestellt werden, dass IT-Sicherheitsbewusstsein einen starken Bezug zu IT-sicherheitsrelevantem Verhalten hat. Dieser Bezug bestimmt auch die Validität der Messung.

Auf Grundlage dieser Erkenntnis wurden die bisher zur IT-Sicherheitsbewusstseinsmessung eingesetzten Methoden – Befragungen, Laborexperimente, Feldexperimente, Sicherheitsmetriken und hybride Ansätze – bewertet. Diese Bewertung zeigt auf, dass die Messmethode der Experimente sich besonders zur Verhaltensforschung eignet. Befragungen eignen sich zur Erfassung von Einflussfaktoren auf den dem Verhalten zugrundeliegenden Entscheidungsprozess. Aus diesen Vorteilen ergibt sich die Motivation hybrider Ansätze zur Messung. Jedoch zeigen sich auch Schwächen bei dem Einsatz bisheriger Experimente. Insbesondere ist die kleine Varianz an Reizgebern und die mangelnde Skalierung zu bemängeln. In Folge der beschränkten Varianz ist auch die Reliabilität dieser Messungen unbekannt.

Die zweite Forschungsfrage adressiert mögliche Verbesserungsmöglichkeiten auf Basis existierender Ansätze. Die herausgearbeiteten Schwächen bisheriger

Methoden motivieren die Konzeption der artefaktbasierten IT-Sicherheitsbewusstseinsmessung als Antwort auf die zweite Forschungsfrage. Diese Form der Messung hat den Vorteil, ein breiteres Spektrum an Reizgebern zulassen zu können.

Die Quantifizierung des IT-Sicherheitsbewusstseins der Probanden als Individuen und als Gruppe durch zwei Maßzahlen auf Basis der im Test ergriffenen Handlungsoptionen erlauben die Skalierung der Messung. Damit wird die dritte Forschungsfrage nach einem konsolidierten Maß für IT-Sicherheitsbewusstsein beantwortet. Um das Konzept zu implementieren und diese Implementierung auch außerhalb des akademischen Rahmens anwendbar zu gestalten, muss die Messung nach diesem Konzept ökonomisch umsetzbar sein. Aus diesem Grund wurde ein Werkzeug gestaltet, das artefaktbasierte IT-Sicherheitsbewusstseinsmessungen unterstützt.

Zur Konzeption dieses Werkzeugs wurden juristische und ethische Rahmenbedingungen erfasst und ausgewertet. Die Beachtung dieser beim Entwurf des Werkzeugs machen einen rechtskonformen und ethischen Einsatz möglich. Dieses Werkzeug wurde im Rahmen eines Feldexperiments erprobt. Die Implementierung des Werkzeugs und die Demonstration seines Einsatzes beantworten die vierte Forschungsfrage nach der ethischen und rechtskonformen Umsetzbarkeit der artefaktbasierten IT-Sicherheitsbewusstseinsmessung. Auf Grundlage der in diesem Experiment erfassten Daten konnte das Konzept der artefaktbasierten IT-Sicherheitsbewusstseinsmessung verifiziert werden.

7.1 Diskussion

Arbeiten zur Messung von IT-sicherheitsrelevantem Verhalten oder zu diesem selbst laden dazu ein, den Nutzer in die Verantwortung zu nehmen. Dies sollte jedoch sehr bedacht geschehen. Die Gestaltung der Informationstechnik und ihre Nutzer sind untrennbar miteinander verknüpft und existieren in steter Wechselwirkung. Aus diesem Grund ist das Wissen um den Nutzer für die adäquate Gestaltung von Technologie von unschätzbarem Wert. Ebenso sind es damit die Werkzeuge, die als Instrumente bei der Gewinnung dieses Wissens zum Einsatz kommen.

Im Rahmen dieser Arbeit wurde nicht nur ein derartiges Instrument weiter entwickelt, die Ergebnisse seiner Erprobung werfen auch Fragen auf. So wird demonstriert, dass eine typische Intervention, deren regelmäßige Durchführung empfohlen ist, auch negative Konsequenzen haben kann. In dem durchgeführten Experiment zeigten die Probanden eine geringere Kontaktaufnahme mit dem Helpdesk, nachdem sie an einer Schulung zur Bildung von IT-Sicherheitsbewusstsein teilnahmen. Organisationen sind auf die Kommunikation ihrer Mitglieder angewiesen. Aber gerade diese Kommunikation ist durch eine Schulung möglicherweise gehemmt. Leider war es nicht möglich, diese Effekte vollständig bestandskräftig nachzuweisen.

Insbesondere zeigt sich, dass der Effekt einer Schulung, mit der die Validität der artekfaktbasierten IT-Sicherheitbewusstseinsmessung geprüft wird, nicht signifikant ist. *P*-Werte unterliegen jedoch einer hohen Variabilität unter Replikation von Studien [LFC12] und haben in diesem Bereich einen schwindenden Einfluss auf das Vertrauen von Wissenschaftlern in das Ergebnis [PL01]. An dieser Stelle ergeben sich mehrere mögliche Erklärungen. Denkbar ist, dass die Intervention keinen Effekt auf die artefaktbasierte Messung von IT-Sicherheitsbewusstsein zeigt. Entweder, weil die Messmethode entgegen der Erwartung ungeeignet ist oder weil die Schulung keinen großen Effekt auf das IT-Sicherheitsbewusstsein hat. Dagegenzuhalten ist jedoch, dass sich dieser Effekt nicht unter den Probanden im Pretest, vor der Intervention, feststellen lässt. Hier kann der Vergleich mit einer etablierten Messmethode weiteren Aufschluss bringen. Diese Daten liegen jedoch entgegen entsprechender Planung nicht vor (vgl. dazu Abschnitt 2.3 und Abschnitt 6.1.3).

Eine weitere Option ist, dass der Interventionseffekt mit der Zeit abnimmt und der Messzeitpunkt zu spät gewählt wurde, um sich in der Messung zu zeigen. Zwischen der Intervention und der zweiten Messphase liegen sieben Monate. Das kann sich negativ auf die Messbarkeit des gewünschten Effekts auswirken [HAO11]. Dieser Umstand hat sich aus dem koordinatorischen Aufwand des Feldexperiments ergeben. Es eröffnet sich Raum für weitere interdisziplinäre Forschung. Neben einer Replikation des Experiments zur Validation kann ein weiterer Durchlauf mit qualitativen Forschungsmethoden begleitet werden, um mögliche Erklärungen für die beobachteten Effekte zu finden.

Auch stellt sich die Frage nach der Bewertung eines Artefakt-Testzeit-Verhältnisses für derartige Messungen. Dieses war im Rahmen des durchgeführten Experiments zwangsweise so gewählt, dass viele Artefakte auf einen kurzen Zeitzeitraum verteilt wurden. Dies kann zu unerwünschten Effekten führen, die sogar die Ergebnisse beeinflussen könnten. Hier gilt es Mindestanforderungen herzuleiten.

Die Ergebnisse der Studie haben jedoch auch einen anleitenden Charakter. Zum einen ist festzuhalten, dass die gewonnenen Verhaltensdaten stark von den Artefakten abhängen. Da für die Testphasen nicht die gleichen Artefakte zu nutzen sind, sind Pretest-Posttest-Studien als nicht aussagefähig anzunehmen. Dieses Ergebnis ist auch auf Phishing-Experimente übertragbar. Zum anderen sind die Messungen mittels Artefakten hinreichend reliabel. Es ist davon auszugehen, dass eine aussagekräftige Messung schon mit wenigen Artefakten erreicht werden kann. An dieser Stelle bedarf es der Schaffung weiterer Erfahrungswerte.

7.2 Ausblick

Die bedeutendste, sich ergebende Folgefragestellung gilt der Verfikation der Beobachtungen zu den Effekten und der Schulung und deren Ursachen. Damit ist die Durchführung eines weiteren Experiments von besonderem Interesse – vorwiegend, um die Beobachtungen bezüglich des Interventionseffekts zu sichern und zu erklären. Auch sollte das Experiment über einen längeren Zeitraum stattfinden, um ein angemessenes Artefakt-Testzeit-Verhältnis gewährleisten zu können und den Erhalt der Interventionseffekte über die Zeit zu studieren.

Darüber hinaus ist der Ausbau der artefaktbasierten Messung von IT-Sicherheitsbewusstsein zu einer hybriden Methode erstrebenswert. Ein möglicher Einsatz ist die Verbindung mit etablierten Werkzeugen wie dem HAIS-Q-Fragebogen [Par+13a; Par+14; McC+16; Par+17]. Hier stellt sich jedoch die Frage, wie eine umfassende Rücklaufquote gesichert werden kann. Über diese Methode können auch demografische Daten über die Probanden erfasst werden. Der Ausbau der artefaktbasierten Messung von IT-Sicherheitsbewusstsein zu einer hybriden Methode stellt einen ersten Ansatz dar, die Beobachtungen erklären zu können.

In Zukunft ist auch das Werkzeug weiterzuentwickeln. Der im Rahmen dieser Arbeit genutzte Entwicklungsstand des Werkzeugs ermöglicht nicht festzustellen, ob ein Artefakt für den Probanden sichtbar war, wenn dieser nicht auf das Artefakt reagiert hat. Damit ist nicht zu unterscheiden, ob der Proband das Artefakt wahrnahm, aber nicht darauf reagierte, oder es erst gar nicht wahrgenommen hat. Das Werkzeug wurde unter der MIT-Lizenz veröffentlicht [SW21]. Es steht damit Forschern und interessierten Anwendern für weitere Untersuchungen, z. B. zur Bewertung unterschiedlicher Interventionen, zur Verfügung.

Darüber hinaus bedarf es auch einer Gütebewertung für Artefakte. Wenn sich Artefakte zueinander in Relation setzen lassen, dann können die Messungen miteinander in Bezug gesetzt werden. Damit würden sich auch Pretest-Posttest-Studiendesigns umsetzen lassen. Erste Ansätze für Phishing-Tests zeigen sich vielversprechend [SGT19].

Auch zeigen die im Rahmen dieser Arbeit hergeleiteten Metriken Potenzial für weitere Forschung. Insbesondere bei der Definition für die Metrik des individuellen IT-Sicherheitsbewusstseins wurde ein konservativer Ansatz gewählt, der mit bisherigen akademischen Publikationen kompatibel ist. Der progressive Ansatz der Herleitung einer Metrik für das IT-Sicherheitsbewusstsein von Gruppen hat neue Einblicke in die Wirkung von Schulungen erlaubt. Mit verbesserter Beobachtung der für das Experiment relevanten Ereignisse eröffnen sich neue Möglichkeiten für deren Interpretation.

Auch zu beachten sind mögliche Nebeneffekte auf die Probanden durch den wiederholten oder kontinuierlichen Einsatz dieser Methode. Werden die Probanden über einen längeren Zeitraum Artefakten als Reizgebern ausgesetzt, ist möglich, dass sich dies auf das Verhalten der Probanden auswirkt. Dies gilt es zu erforschen, bevor die artefaktbasierte Messung von IT-Sicherheitsbewusstsein bedenkenlos in Standards empfohlen werden kann.

Appendix

A Artefakte der Studie

In diesem Kapitel sind die in der Studie (siehe Kapitel 6) zum Einsatz gekommenen Artefakte dokumentiert. Jedes Artefakt wird anhand von Bildschirmaufnahmen aus dem Testumfeld visuell dargestellt. Der technische Abschluss der Artefaktpräsentation (vgl. Abschnitt 5.3.3) kann aufgrund der hohen Bandbreite nicht konsistent erfasst werden. Das Ereignis wird zu jedem Artefakt verzeichnet. Artefakte können eine Serie von Interaktionsmöglichkeiten bieten (siehe Abschnitt 3.2). Die durch den Probanden wählbaren Handlungsoptionen in Form von *überwachten Ereignissen* werden dargestellt. Diese werden durch das Werkzeug über das Reaktionsmonitoring (vgl. Abschnitt 5.3.4) aufgenommen. Alle Artefakte erfüllen Bedingung 11 (Seite 62).

Selbstlöschende Datei

Eine ausführbare Datei mit dem Dateinamen `TopSecret.exe` wird auf dem Desktop des Probanden platziert. Ein Symbol ist nicht in der Datei hinterlegt. Damit fällt die Darstellung auf das generische Symbol für eine ausführbare Datei zurück. Ein derartiges Programm wird im Testumfeld nicht eingesetzt und gehört damit auch nicht zur Arbeitsplatzkonfiguration der Probanden. Die technische Artefaktpräsentation ist beendet, wenn die Datei auf dem Desktop platziert wurde.

Startet der Proband die hinterlegte Datei, löscht der Prozess die Datei und beendet sich danach. Es handelt sich damit um ein statisches hostbasiertes Artefakt.

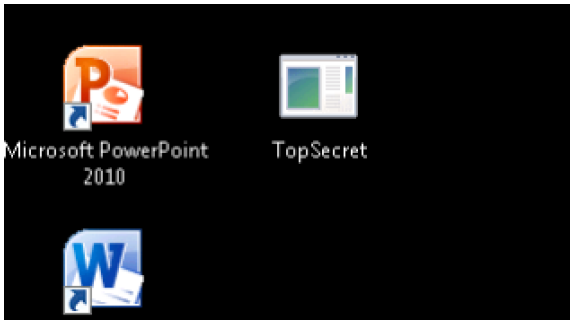


Abbildung 27: *Bildschirmaufnahme: Ausführbare Datei, die sich beim Start selbst löscht.*

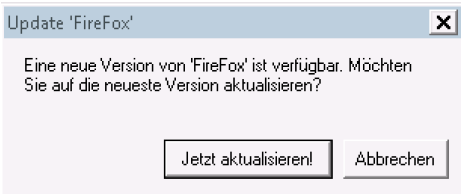
Überwachte Ereignisse

Wird ein überwachtes Ereignis ausgelöst, wird eine HTTP-Anfrage, die den Nutzernamen des eingeloggten Nutzers beinhaltet, an APE gesendet.

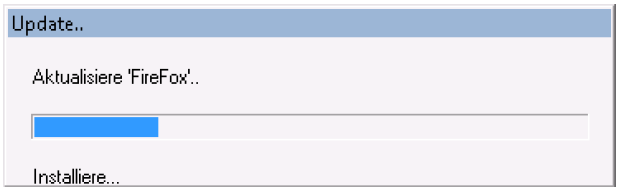
Ausführen Die Ausführung der Datei, z. B. durch einen Doppelklick auf das Programmsymbol, wird registriert.

Updater

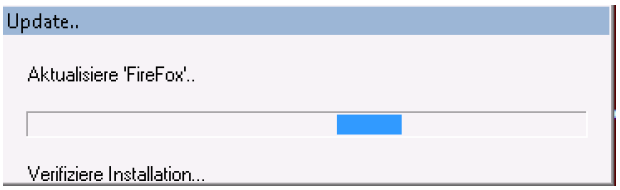
Das *Updater*-Artefakt ist ein dynamisches hostbasiertes Artefakt. Wird es ausgeführt, wird dem Probanden ein Dialog präsentiert. Der Text dieses Dialogs besagt, dass eine neue Version eines Programms namens *FeuerFuchs* existiert. Der Programmname erinnert an den wohlbekannten Browser *Firefox* [Moz16]. Es schließt die Frage an, ob auf die neueste Version dieses Programms aktualisiert werden soll. Ein Programm mit den Namen *FeuerFuchs* wird im Testumfeld nicht eingesetzt und gehört damit auch nicht zur Arbeitsplatzkonfiguration der Probanden (vgl. Abschnitt 6.1.1). Darüber hinaus müssen Aktualisierungen im Testumfeld nicht durch den Nutzer und damit auch nicht durch die Probanden bestätigt werden. Die technische Artefaktpräsentation ist beendet, wenn der Dialog angezeigt wird.



(a) Wird die Aktion „Jetzt aktualisieren!“ gewählt, so wechselt die Ansicht auf eine Fortschrittsanzeige.



(b) Die Fortschrittsanzeige entwickelt sich selbstständig über den Verlauf von zehn Sekunden.



(c) Die anschließende Aktivitätsanzeige bleibt fünf Sekunden lang erhalten und soll dem Probanden die Verifikation der Installation suggerieren.

Abbildung 28: *Bildschirmaufnahme: Betriebssystemdialog, der zur Aktualisierung eines Programms namens* FireFox *auffordert. Diese Abbildung illustriert das Artefakt mit dem Term* FireFox. *In der Testdurchführung kam der Name* FeuerFuchs *zum Einsatz.*

Überwachte Ereignisse

Wird ein überwachtes Ereignis ausgelöst, wird eine HTTP-Anfrage, die den Nutzernamen des eingeloggten Nutzers beinhaltet, an APE gesendet.

Start Der Start des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit der Anzeige des Dialogs auf.

Manipulation Alle Manipulationen der Fenstersteuerungselemente, z. B. das Schließen des Fensters durch das „X“ in der Fensterkante rechts oben sowie die Betätigung der angebotenen Schaltflächen, werden registriert. Wird

die vermeintliche Aktualisierung durch den Probanden angestoßen, wird diese, wie in Abbildung 28 dargestellt, visualisiert.

Ende Das Ende des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit dem Ende des Prozesses auf. Dieses Ereignis reflektiert keine Handlungsoption.

Targo Bank

Dieses Artefakt ist eine Phishing-E-Mail, die an die Teilnehmer der Studie versandt wurde. Sie ist nicht personalisiert und adressiert den Empfänger mit *Kunde*. Der Adressat ist aufgefordert, seine PIN aus Sicherheitsgründen zu ändern. Die E-Mail enthält eine URL und dem Adressaten wird erklärt, dass er seine PIN unter diesem Hyperlink kostenlos ändern kann. Wenn der Adressat der Aufforderung nicht Folge leistet, würde das Konto mit 18,99 € belastet und der Adressat postalisch erneut zur Änderung der PIN aufgefordert werden (vgl. Abbildung 29). Die technische Artefaktpräsentation ist beendet, wenn die E-Mail an den Mailserver übergeben wurde.

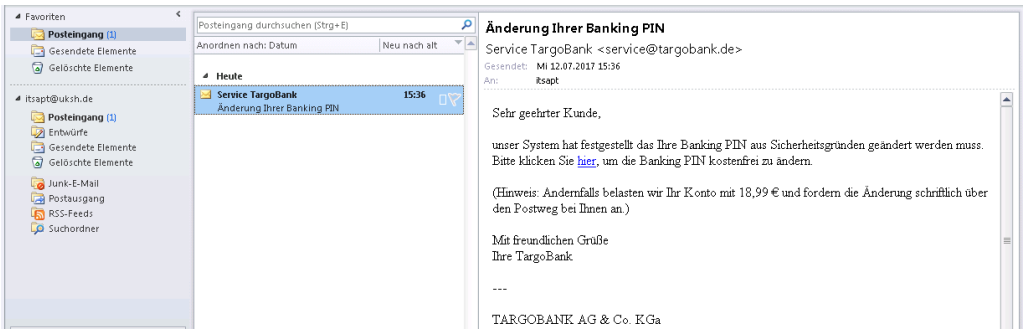


Abbildung 29: *Bildschirmaufnahme: Phishing-E-Mail des Targo-Bank-Artefakts.*

Folgt der Teilnehmer dem angebotenen Link zu `http://tarrgobank.de/change_pin/index.php`, gelangt er auf die in Abbildung 30 dargestellte Webseite. Die Webseite bietet ein Formular an, um einen Vor- und Zunamen sowie eine alte und neue PIN (zwei Mal) einzugeben.

Betätigt der Proband die mit „Abschicken“ ausgezeichnete Schaltfläche, wird er auf die Domäne `http://example.org` umgeleitet (vgl. Abbildung 31).

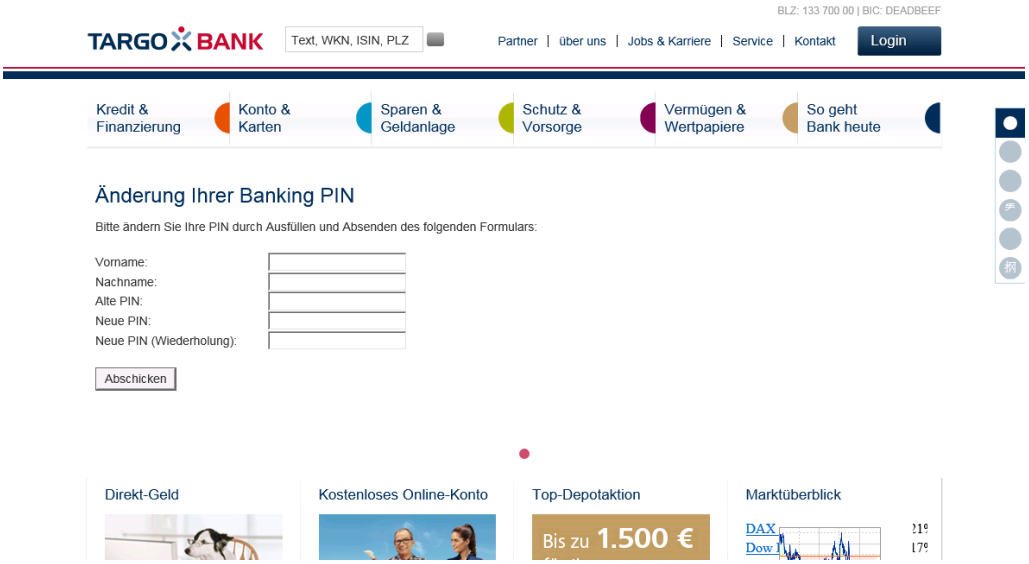


Abbildung 30: *Bildschirmaufnahme: Präparierte Webseite des Targo-Bank-Artefakts.*

Abbildung 31: *Bildschirmaufnahme: Abschluss-Webseite des Targo-Bank-Artefakts.*

Überwachte Ereignisse

Der Besuch der Webseite sowie das Senden der Formulardaten werden über das Log des HTTP-Servers erfasst und an APE gesendet.

Besuch Der Besuch der Webseite wird überwacht.

Abschicken Betätigt der Proband die mit „Abschicken“ ausgezeichnete Schaltfläche, wird diese Aktion aufgezeichnet. Die an den Server übermittelten Daten werden nicht aufgezeichnet.

Nur Link

Dieses Artefakt ist eine Phishing-E-Mail, die lediglich einen Hyperlink zu einer durch einen verbreiteten Kurz-URL-Dienst [Sch10] verdeckten URL beinhaltet (vgl. Abbildung 32). Die E-Mail trägt als Betreff den Namen des Kurz-URL-Dienstes, *bit.ly*. Die technische Artefaktpräsentation ist beendet, wenn die E-Mail an den Mailserver übergeben wurde.

Abbildung 32: *Bildschirmaufnahme: E-Mail des Link-Only-Artefakts.*

Wird die verknüpfte URL aufgerufen, wird eine Webseite ausgeliefert, die den HTTP-Fehler 404 anzeigt (vgl. Abbildung 33).

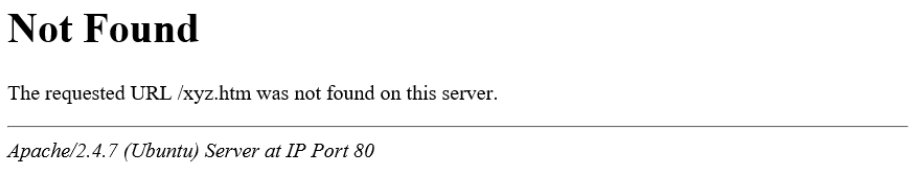


Abbildung 33: *Bildschirmaufnahme: Webseite des Link-Only-Artefakts.*

Überwachte Ereignisse

Der Besuch der Webseite wird über das Log des HTTP-Servers erfasst und von APE gesendet.

Besuch Der Besuch der Webseite wird überwacht.

IT Ticket

Das IT-Ticket-Artefakt ist eine Phishing-E-Mail, die einer automatisch generierten E-Mail des *IT Service Desk* nachempfunden ist. Der in der E-Mail verwendete Text mischt die Sprachen Deutsch und Englisch. Der eingebettete

Hyperlink beinhaltet `uksh`, die offizielle Abkürzung des Arbeitgebers, als *fourth level domain*. Die bezeichnende *second level domain* ist `it-ticket`. Die Zieldomain ist keine legitime Domain des Arbeitgebers und auch nicht als Domain eines Geschäftspartners bekannt. Die technische Artefaktpräsentation ist beendet, wenn die E-Mail an den Mailserver übergeben wurde.

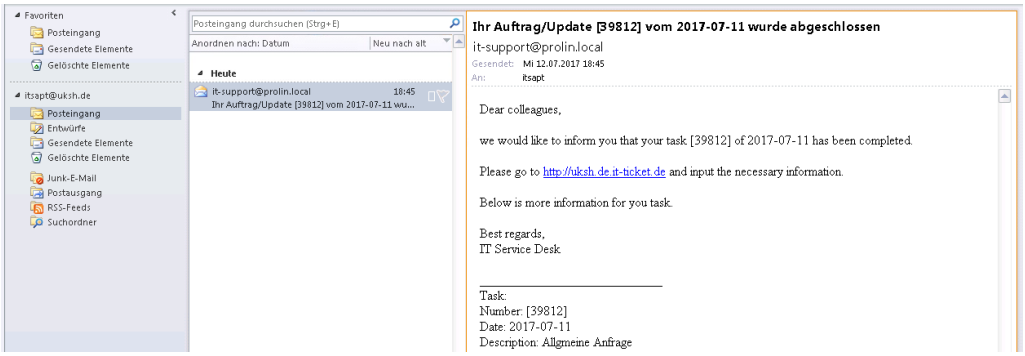


Abbildung 34: *Bildschirmaufnahme: E-Mail des IT-Ticket-Artefakts.*

Folgt der Proband dem Hyperlink, wird ein rudimentäres Formular präsentiert, das dem Internetauftritt des Arbeitgebers ähnelt (aber nicht gleicht). Es bietet die Formularfelder *Vorname*, *Zuname*, *(Auftrags-) Nummer*, *(Auftrags-) Datum* und ein größeres Textfeld für die Eingabe einer Nachricht an.

Versendet der Proband das Formular, wird er wie beim Targo-Bank-Artefakt (vgl. Abbildung 31) auf die Domain `http://example.org` umgeleitet.

Überwachte Ereignisse

Der Besuch der Webseite sowie das Senden der Formulardaten werden über das Log des HTTP-Servers erfasst und an APE gesendet.

Besuch Der Besuch der Webseite wird überwacht.

Senden Sendet der Proband Daten an den Server, wird diese Aktion aufgezeichnet. Die an den Server übermittelten Daten werden nicht aufgezeichnet.

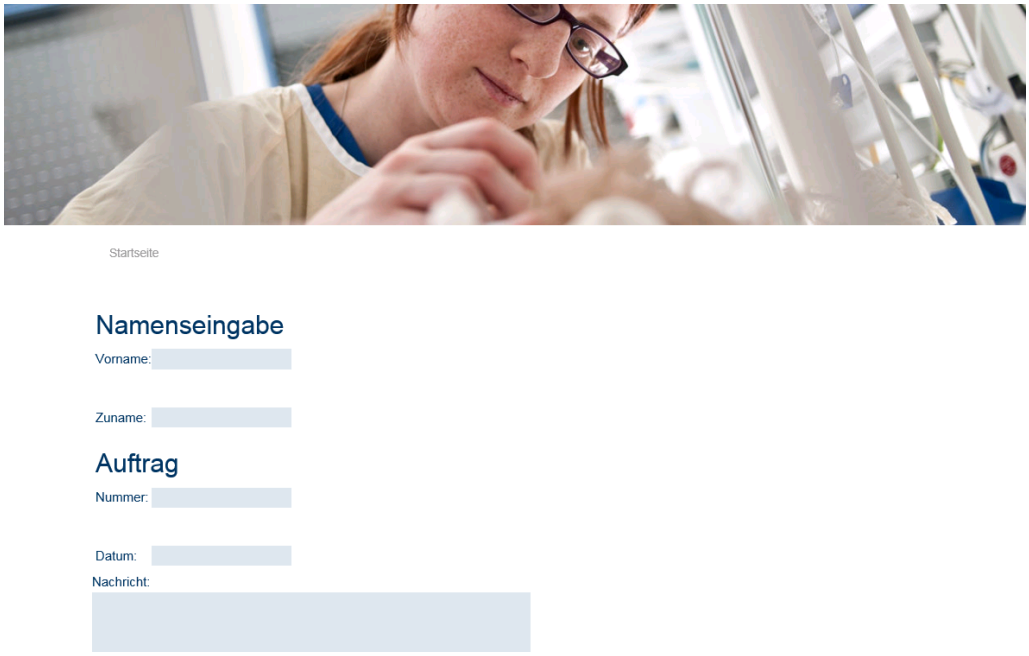


Abbildung 35: *Bildschirmaufnahme: Webseite des IT-Ticket-Artefakts.*

Anti Virus

Das Artefakt *Anti Virus* ist ein Betriebssystemdialog. Wenn der Test startet, wird der Dialog präsentiert. Der Dialog verschwindet, wenn eine Schaltfläche oder ein Fenstersteuerungselement betätigt wird. Die technische Artefaktpräsentation ist beendet, wenn der Dialog angezeigt wird.

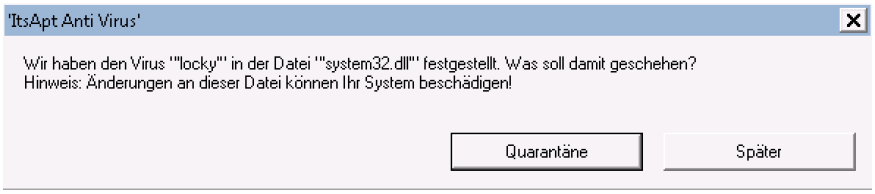


Abbildung 36: *Bildschirmaufnahme: Dialog des Anti-Virus-Artefakts. Während der Testdurchführung wurden abweichende Zeichenketten genutzt. Die Titelleiste verzeichnete „AntiVirus“, der Virus wurde „virus.exe“ und die Datei wurde „vscr32.dll“ genannt.*

Überwachte Ereignisse

Wird ein überwachtes Ereignis ausgelöst, wird eine HTTP-Anfrage, die den Nutzernamen des eingeloggten Nutzers beinhaltet, an APE gesendet.

Start Der Start des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit der Anzeige des Dialogs auf.

Manipulation Alle Manipulationen der Fenstersteuerungselemente, z. B. das Schließen des Fensters durch das „X“ in der Fensterkante rechts oben sowie die Betätigung der angebotenen Schaltflächen, werden registriert.

Ende Das Ende des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit dem Ende des Prozesses auf. Dieses Ereignis reflektiert keine Handlungsoption.

Login-Fenster

Das Artefakt *Login-Fenster* ist ein Betriebssystemdialog. Es bietet zwei Textfelder, die mit „Benutzername“ und „Kennwort“ beschriftet sind. Letzteres reflektiert die Eingabe in Form eines Sternsymbols „*“ für jedes eingegebene Zeichen. Es werden zwei Schaltflächen, beschriftet mit „OK“ und „Abbrechen“, angeboten. Die technische Artefaktpräsentation ist beendet, wenn der Dialog angezeigt wird.

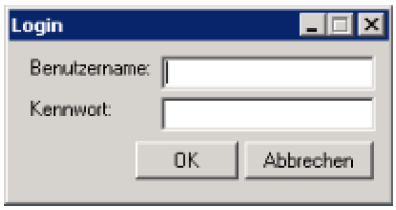


Abbildung 37: *Bildschirmaufnahme: Dialog des Login-Fenster-Artefakts.*

Überwachte Ereignisse

Wird ein überwachtes Ereignis ausgelöst, wird eine HTTP-Anfrage, die den Nutzernamen des eingeloggten Nutzers beinhaltet, an APE gesendet. Die eventuell durch den Nutzer eingegebenen Daten werden nicht überwacht, sondern direkt verworfen.

Start Der Start des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit der Anzeige des Dialogs auf.

Manipulation Alle Manipulationen der Fenstersteuerungselemente, z. B. das Minimieren oder Schließen des Fensters durch das „X“ in der Fensterkante rechts oben sowie die Betätigung der angebotenen Schaltflächen, werden registriert.

Ende Das Ende des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit dem Ende des Prozesses auf. Dieses Ereignis reflektiert keine Handlungsoption.

Defacing

Um das Defacing-Artefakt präsentiert zu bekommen, muss der Proband das Intranet des Arbeitgebers öffnen. Alle Bilder, die mithilfe eines `img`-Tags im HTML-Dokument eingebunden werden, erscheinen um 180° gedreht. Das Bild im Banner in der Mitte der Seite (vgl. Abbildung 38) ist mittels der CSS-Anweisung `background-image` (*Cascading Style Sheets* [W3C18]) mit dem entsprechenden Element verknüpft und deshalb nicht von der Manipulation betroffen.

Das Intranet ist als Startseite des Browsers in der Auslieferungskonfiguration der Bildschirmarbeitsplätze gesetzt. Hat der Proband die Startseite seines Browsers nicht verändert, wird das Artefakt automatisch beim Start des Browsers präsentiert.

Gemäß dem Grundsatz der Vermeidung von Nicht-Verfügbarkeit (vgl. Abschnitt 3.2) ist die Webseite voll funktionsfähig. Wird ein Hyperlink aktiviert oder die Webseite neu geladen, erscheint sie wieder unmodifiziert. Die technische Artefaktpräsentation ist beendet, wenn die Webseite abrufbar ist.

Überwachte Ereignisse

Der Besuch der Webseite sowie jedes Aktivieren eines Hyperlinks werden über das Log des HTTP-Servers erfasst und an APE gesendet.

Besuch Der Besuch der Webseite wird überwacht.

Abbildung 38: *Bildschirmaufnahme: Webseite des Defacing-Artefakts.*

Interaktion Aktiviert der Proband einen Hyperlink, wird diese Aktion aufgezeichnet. Eventuell an den Server übermittelte Daten werden nicht aufgezeichnet.

IT Department

Bei diesem Artefakt handelt es sich um eine Phishing-E-Mail. Die an die Probanden verschickte E-Mail erklärt, es hätte einen Ausfall im Rechenzentrum gegeben. Der Empfänger wird aufgefordert, sich unter dem Link `http://uksh.de.owauksh.de` mit seinem Nutzernamen und Passwort anzumelden und die Funktionsfähigkeit dieses Dienstes zu prüfen.

Die Domain `owauksh.de` beinhaltet sowohl die offizielle Abkürzung des Arbeitgebers als auch die Buchstaben des Akronyms für den im Unternehmen wohl bekannten Dienst *Outlook Web App*. Darüber hinaus ist die angegebene Domain weder bekannt noch den Probanden in diesem Kontext vertraut. Die technische Artefaktpräsentation ist beendet, wenn die E-Mail an den Mailserver übergeben wurde.

Folgt der Proband dem angebotenen Link, wird er mit einem originalgetreuen Nachbau der Login-Seite der *Outlook Web App* konfrontiert. Diese wird in Abbildung 40 dargestellt.

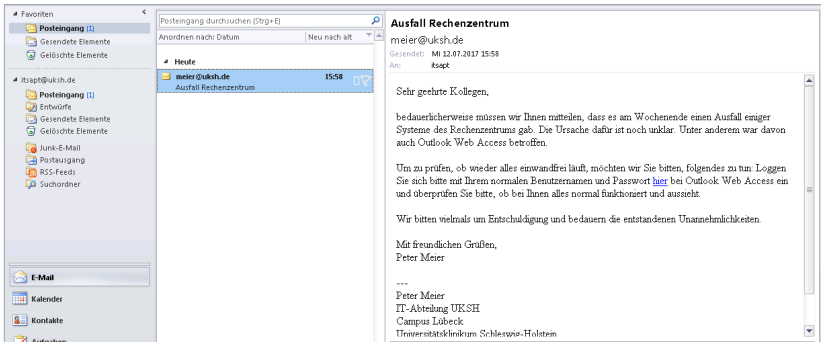


Abbildung 39: *Bildschirmaufnahme: E-Mail des IT-Department-Artefakts.*

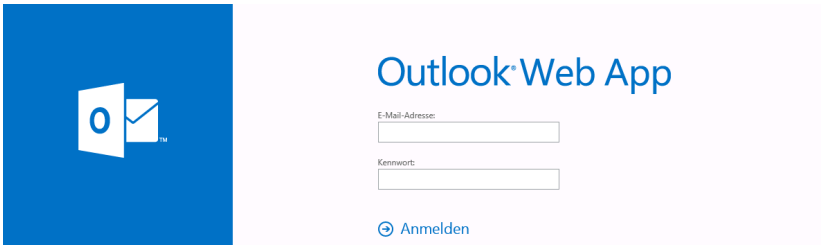


Abbildung 40: *Bildschirmaufnahme: Webseite des IT-Department-Artefakts.*

Aktiviert der Proband den mit „Anmelden“ beschrifteten Hyperlink, wird er wie beim Artefakt Anhang auf die Domain `http://example.org` umgeleitet.

Überwachte Ereignisse

Der Besuch der Webseite sowie das Senden der Formulardaten werden über das Log des HTTP-Servers erfasst und an APE gesendet.

Besuch Der Besuch der Webseite wird überwacht.

Abschicken Sendet der Proband Daten an den Server, wird diese Aktion aufgezeichnet. Die an den Server übermittelten Daten werden nicht aufgezeichnet.

Versichertenkarte

Dieses Artefakt ist eine Phishing-E-Mail. Der Text der E-Mail wendet sich an eine nicht weiter spezifizierte Ansprechperson. Der Autor würde hiermit

seine „Versichertenkarte“ übermitteln. Die Signatur dieser E-Mail suggeriert, dass diese E-Mail auf einem Mobiltelefon verfasst wurde. Die technische Artefaktpräsentation ist beendet, wenn die E-Mail an den Mailserver übergeben wurde.

Abbildung 41: *Bildschirmaufnahme: E-Mail des Versichertenkarte-Artefakts.*

An der E-Mail befindet sich ein Anhang mit dem Dateinamen `Versichertenkarte.html`. Versucht der Proband die Outlook-Funktion *Vorschau* zu nutzen, um den E-Mail-Anhang zu inspizieren, erhält er eine Warnung (vgl. Abbildung 42).

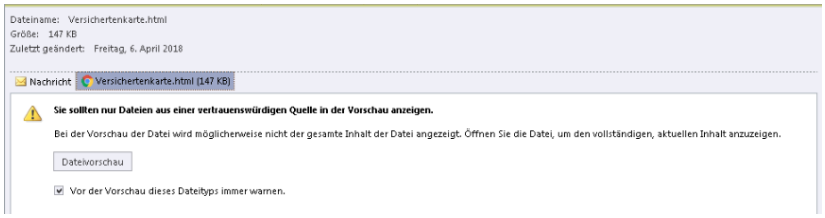


Abbildung 42: *Bildschirmaufnahme: Warnung des Versichertenkarte-Artefakts.*

Aktiviert der Proband die Dateivorschau trotz der vorhergehenden Warnung, wird ein Platzhalter mit dem Text „Gesundheitskarte.jpg“ angezeigt. Darüber hinaus erklärt ein Banner, dass Bilder nur angezeigt werden können, wenn das Element geöffnet wird (vgl. Abbildung 43)

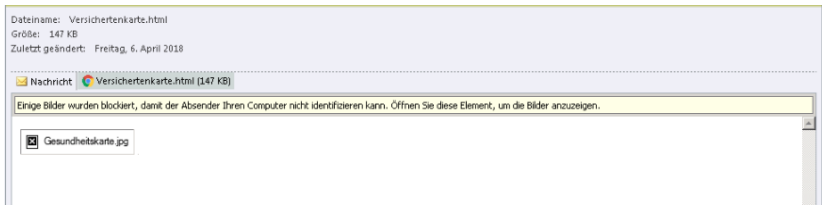


Abbildung 43: *Bildschirmaufnahme: Vorschau des Versichertenkarte-Artefakts.*

Öffnet der Proband den Dateianhang, zeigt sich die Gesundheitskarte im Browser.

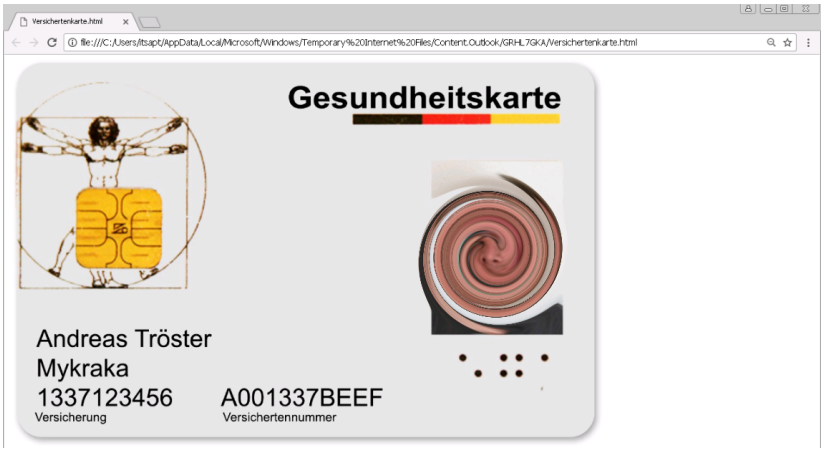


Abbildung 44: *Bildschirmaufnahme: Dateianhang des Versichertenkarte-Artefakts. Die Person in der Bildschirmaufnahme ist aus Datenschutzgründen in dieser Arbeit unkenntlich gemacht. Der Name, die Versicherungsnummer und die Versicherungskennnummer sind frei erfunden.*

Überwachte Ereignisse

Beim Öffnen des Anhangs wird eine HTTP-Anfrage an APE gesendet.

Anhang Das Öffnen des Anhangs wird überwacht.

Java Update

Das Java-Update-Artefakt ist ein Betriebssystemdialog. Der Text besagt, dass ein Update für Java™ verfügbar sei. Es werden zwei Schaltflächen mit den Beschriftungen „Installieren“ und „Später“ angeboten. Die technische Artefaktpräsentation ist beendet, wenn der Dialog angezeigt wird.

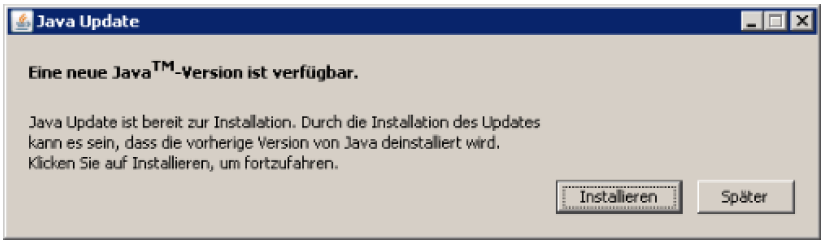


Abbildung 45: *Bildschirmaufnahme: Betriebssystemdialog des Java-Update-Artefakts.*

Wird die Schaltfläche mit der Beschriftung „Später“ betätigt, schließt sich der Dialog. Wird die Schaltfläche mit der Beschriftung „Installieren“ betätigt, wandelt sich der Dialog und es wird eine Fortschrittsanzeige visualisiert. Diese ist mit „Downloading …“ beschriftet. Ist sie vollständig beginnt eine zweite Fortschrittsanzeige mit der Beschriftung „Installing …“.

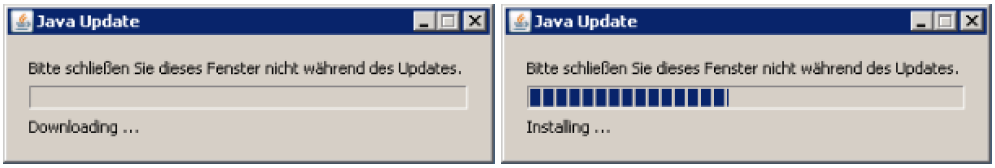


Abbildung 46: *Bildschirmaufnahme: Fortschrittsanzeigen des Java-Update-Artefakts.*

Überwachte Ereignisse

Wird ein überwachtes Ereignis ausgelöst, wird eine HTTP-Anfrage, die den Nutzernamen des eingeloggten Nutzers beinhaltet, an APE gesendet.

Start Der Start des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit der Anzeige des Dialogs auf.

Manipulation Alle Manipulationen der Fenstersteuerungselemente, z. B. das Minimieren oder Schließen des Fensters durch das „X“ in der Fensterkante rechts oben sowie die Betätigung der angebotenen Schaltflächen, werden registriert.

Ende Das Ende des Prozesses wird überwacht. Das Ereignis tritt gemeinsam mit dem Ende des Prozesses auf. Dieses Ereignis reflektiert keine Handlungsoption.

Word Macro

Ein MS Word™-Dokument mit dem Titel „Unbenannt(1).doc“ wird auf dem Desktop des Probanden platziert. Öffnet der Proband das Dokument, präsentiert sich ein gelbes Banner mit der Aufschrift „Sicherheitswarnung“ und gibt weiter an, dass die Makros deaktiviert wurden. Es bietet eine Schaltfläche an, um

den „Inhalt [zu] aktivieren“. Als Inhalt des Dokuments wird ein Pfeil auf den Banner dargestellt und auf englischer Sprache darum gebeten, die Makros dieses Dokuments zu aktivieren, um das Dokument ansehen zu können. Die technische Artefaktpräsentation ist beendet, wenn die Datei auf dem Desktop platziert wurde.

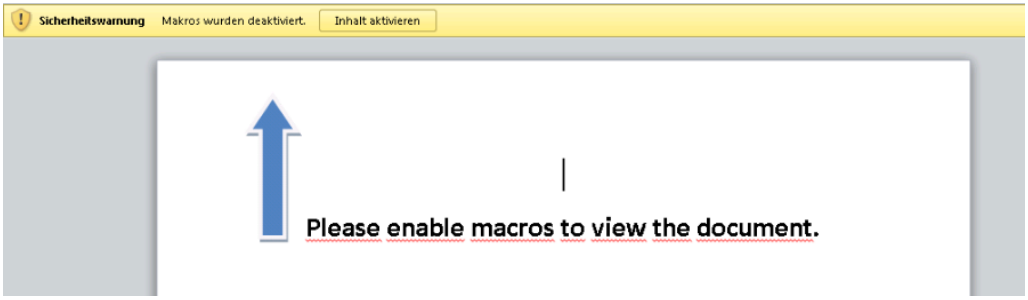


Abbildung 47: *Bildschirmaufnahme: Warnung des Word-Macro-Artefakts.*

Folgt der Proband der Aufforderung, verschwindet das gelbe Banner. Der Inhalt des Dokuments ändert sich jedoch nicht.

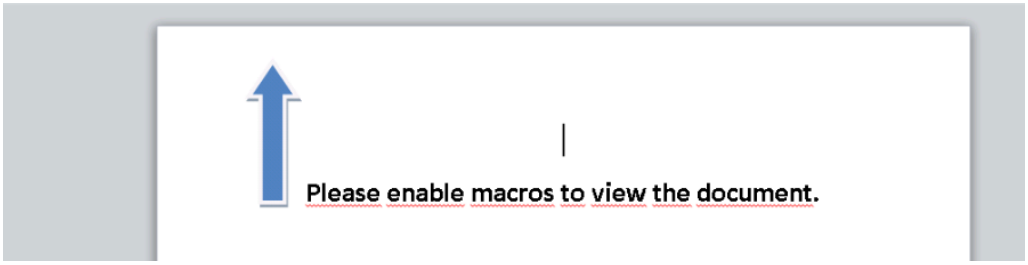


Abbildung 48: *Bildschirmaufnahme: Finale Ansicht des Word-Macro-Artefakts.*

Überwachte Ereignisse

Wird ein überwachtes Ereignis ausgelöst, wird eine HTTP-Anfrage, die den Nutzernamen des eingeloggten Nutzers beinhaltet, an APE gesendet.

Makros Aktiviert der Nutzer die Makros wird ein überwachtes Ereignis ausgelöst.

Manipulation Sollte sich die Datei zum Ende der Testperiode nicht mehr auf dem Desktop befinden, wird ein überwachtes Ereignis ausgelöst.

Bounce

Eine E-Mail wird an die Probanden versendet, die einer E-Mail gleicht, welche von dem eingesetzten Mailserver versendet wird, wenn ein Nutzer versucht, eine E-Mail an einen unbekannten Adressaten zu versenden. Eine derartige E-Mail wird auch *Bounce Notification* genannt [GK11]. Die technische Artefaktpräsentation ist beendet, wenn die E-Mail an den Mailserver übergeben wurde.

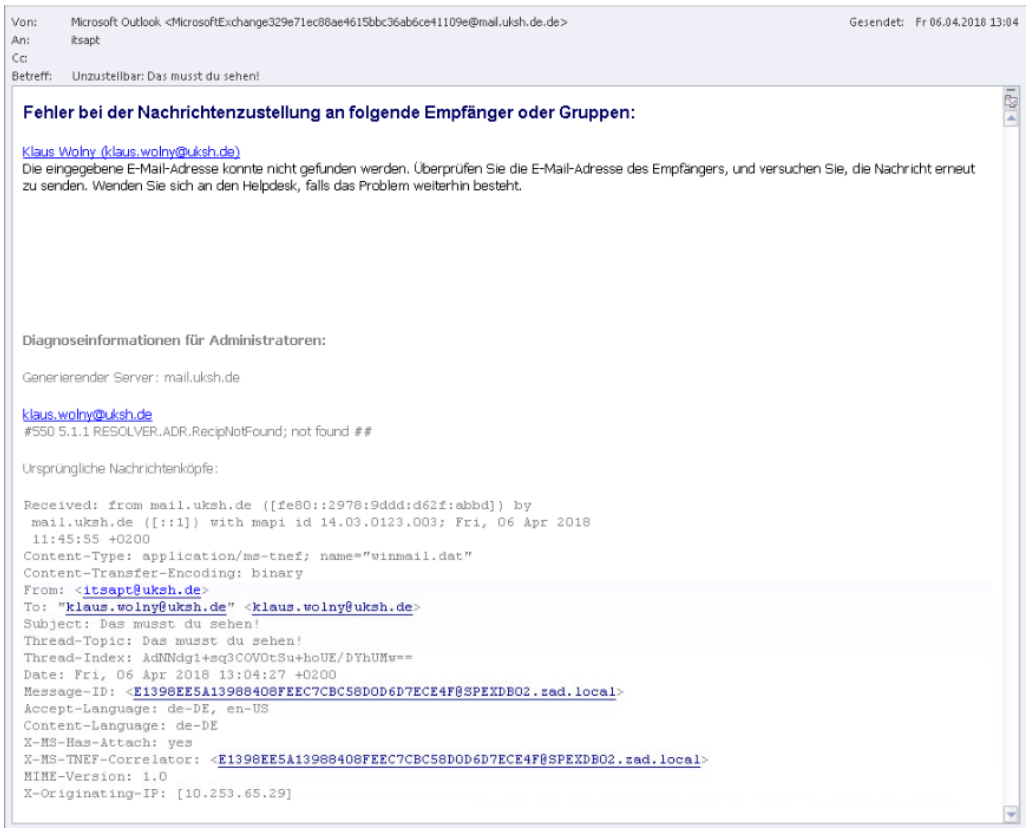


Abbildung 49: *Bildschirmaufnahme: E-Mail des Bounce-Artefakts.*

Das Bounce-Artefakt bietet dem Nutzer keine Handlungsoption.

B Teilnahmedaten der Probanden

Die Tabellen 11 und 12 enthalten die Teilnahmedaten der Probanden für die Elemente des Experiments in anonymisierter Form. Eine Zeile spiegelt einen Probanden wieder. Probanden, die einer Nutzung ihrer Daten widersprochen haben (vlg. Abschnitt 6.2), sind exkludiert. Die *Motivation* eines Probanden an der Schulung teilzunehmen, ist mit „1“ für seine bekundete Motivation und mit „0“ für das Nicht-bekunden codiert. Die Teilnahme an der *Schulung* ist mit „1“ codiert. Hat der Proband nicht an der Schulung teilgenommen, verzeichnet die Spalte den Wert „0“. Alle weiteren Spalten enthalten die aufgezeichneten Reaktionen der Probanden auf die Artefakte. Tabelle 10 verzeichnet das Codierungsschema.

Tabelle 10: *Codierungsschema für die aufgezeichneten Reaktionen (nach Tabelle 2) der Probanden.*

Reaktion	Wert
$\neg i \wedge \neg r$	0
$i \wedge \neg r$	1
$i \wedge r$	2
$\neg i \wedge r$	3
keine Teilnahme	-

Tabelle 11: *Teilnahmen der Probanden mit ID < 197.*

Motivation	Schulung	Anti Virus	Defacing	Selbstlöschende Datei	Updater	Bounce	Java Update	Login-Fenster	Word Macro	IT Ticket	Nur Link	IT Department	Versichertenkarte	Anti Virus
0	0	1	0	0	1	-	1	-	0	0	-	0	0	1
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	0	1	-	-	-	-	0	1
0	0	-	-	-	-	0	-	-	-	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	1	0	-	-	-	-	-	-	0	0	0	0	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	1
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	1	-	0	-	-	-	-	-	0	0	0	0	-
0	0	-	-	-	-	-	1	1	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	-	-	-	-	-	0
0	0	1	1	0	1	-	-	1	0	0	-	0	-	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	3	1	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	0	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	1	-	0	1	-	-	-	-	-	0	0	0	-
0	0	-	-	-	-	0	1	-	-	-	-	-	-	-

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 11.

0	0	1	0	0	-	0	-	1	0	0	0	0	0	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	-
0	0	-	0	-	1	-	-	1	0	0	0	-	0	1
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	1	0	3	1	-	1	1	0	0	0	0	0	-
0	0	-	-	-	-	0	-	1	0	0	-	-	0	1
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	1	1	-	-	-	-	-	0
0	0	1	0	0	-	-	-	-	-	3	0	0	0	-
0	0	1	0	0	1	-	-	1	-	0	0	0	0	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	1	0	1	0	-	-	0	0	3	3	3	0
0	0	0	1	0	-	0	-	-	0	0	0	0	0	0
0	0	-	-	-	-	0	-	-	1	-	-	-	-	-
0	0	-	0	-	1	0	-	-	0	-	0	-	-	1
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	0	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	3	2	0	0	-	1	1	0	0	0	0	0	-
0	0	-	0	0	3	0	-	-	0	0	0	-	3	2
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	0	-	-	-	0	-	-	-	0	-	-	0	0
0	0	0	-	0	0	3	-	1	0	-	0	0	3	0
0	0	-	-	-	-	0	-	-	-	-	-	-	-	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	3

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 11.

0	0	0	1	1	0	0	-	1	0	0	0	0	0	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	0	0	0	0	-	-	-	-	-	-	0	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	1	-	0	0	-	1	0	-	0	-	0	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	1	-	-	-	-	-	0
0	0	0	0	1	0	-	-	1	1	0	0	0	0	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	0	0	0	0	-	-	-	-	0	0	0	0	-
0	0	3	0	0	0	-	1	-	1	0	0	0	0	2
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	0	0	0	0	-	-	-	-	0	0	0	0	-
0	0	-	-	0	3	0	-	1	0	-	0	-	0	3
0	0	0	0	-	0	-	-	-	-	3	0	-	3	-
0	0	0	-	-	-	-	-	-	0	-	0	-	0	1
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	3	0	0	0	-	-	-	-	0	0	0	0	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	1	-	-	-	-	-	-	0
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	1	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	-	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 11.

0	0	-	1	0	0	-	1	-	0	0	0	-	0	3
0	0	0	0	0	0	-	-	-	-	3	0	3	0	-
0	0	-	-	-	-	-	-	1	0	-	-	-	-	1
0	0	-	1	0	0	-	-	-	-	-	0	0	-	-
0	0	-	-	-	-	-	1	1	0	-	-	-	-	-
0	0	-	-	-	-	3	-	-	0	-	-	-	-	3
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	-
0	0	-	-	-	0	-	-	-	-	0	0	-	0	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	2
1	0	-	-	-	-	0	-	-	0	-	-	-	-	-
1	0	-	-	-	-	-	-	-	0	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	0	1	-	0	-	-	-	-	-
1	0	-	-	-	-	0	-	-	0	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	0	-	-	0	-	-	-	-	-
1	0	-	-	-	-	0	-	0	-	-	-	-	-	1
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	0	1	1	-	-	-	-	-	0
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	0	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	0	-	-	-	-	0
1	0	-	-	-	-	0	-	-	0	-	-	-	-	0
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	0	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	1	0	1	0	1	1	0	3	-	-	0	3

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 11.

1	0	-	0	-	1	-	-	-	-	3	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	3	0	0	0	-	1	-	0	0	0	0	0	1
1	0	-	-	0	0	0	1	-	0	-	1	-	0	-
1	0	0	0	0	0	0	-	-	0	0	0	0	0	-
1	0	-	-	-	-	0	-	-	0	-	-	-	-	0
1	0	0	0	-	-	0	1	-	0	0	-	0	0	1
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	0	0	0	0	-	-	-	-	0	0	0	0	-
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	-	-	0	-	-	-	-	0
1	0	-	-	-	-	-	-	-	-	-	-	-	-	-
1	0	-	-	-	0	-	-	-	-	-	-	-	-	-
1	0	-	-	-	-	-	1	1	1	-	-	-	-	1
1	0	0	0	0	0	0	-	-	0	0	0	-	0	-
1	0	0	0	0	0	0	-	-	0	0	0	0	0	0
1	1	-	-	-	-	0	-	-	-	-	-	-	-	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	1	-	-	1	-	-	-	-	-	-	3	-	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	1	-	-	1	-	-	-	-	0	0	0	0	-
1	1	2	1	-	-	-	-	-	-	1	-	3	-	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	-	-	-	-	1	-	-	-	-	-	-	0
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	0	0	1	-	-	1	-	-	0	0	-	2
1	1	-	-	-	-	-	-	-	0	-	-	-	-	1
1	1	2	0	-	-	-	-	-	-	3	-	0	-	-

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 11.

1	1	1	0	-	1	-	-	-	-	0	0	0	0	-
1	1	2	0	0	1	-	-	-	-	0	0	3	0	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	-	-	-	-	-	1	1	1	-	-	-	-	-
1	1	-	-	-	-	0	1	1	0	-	-	-	-	-
1	1	0	0	-	-	-	-	-	-	0	-	0	-	-
1	1	0	0	0	1	-	-	-	-	0	0	3	0	-
1	1	0	-	-	1	-	-	-	-	0	-	0	0	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	-
1	1	0	0	-	1	-	-	-	-	0	0	0	-	-
1	1	3	0	1	0	-	-	-	-	0	0	0	0	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	0
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	-	-	-	-	-	-	-	0	-	-	-	-	-
1	1	-	0	0	0	-	-	-	-	-	0	-	-	-
1	1	-	-	0	-	-	-	-	-	-	0	-	0	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
1	1	0	1	0	0	-	-	-	-	0	0	0	0	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	0	-	0	-	-	-	-	0	0	-	-	0	0
1	1	0	0	0	0	-	-	-	-	0	0	0	0	-
1	1	0	0	0	0	-	-	-	-	0	0	0	0	-
1	1	-	-	-	-	0	-	-	0	-	-	-	-	0
1	1	-	-	-	-	0	-	-	0	-	-	-	-	-
1	1	0	0	0	0	-	-	-	-	0	0	0	0	-
1	1	-	0	0	0	-	-	-	-	0	0	-	0	-
1	1	0	0	-	-	-	-	-	-	3	0	-	3	-
1	1	0	0	-	0	-	-	-	-	0	0	0	0	-
1	1	0	-	-	-	-	-	-	-	-	-	0	0	-
1	1	-	-	-	-	-	-	-	-	-	-	-	-	-

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 11.

1	1	-	-	-	-	-	-	-	0	-	-	-	-	0
1	1	0	0	-	0	-	-	-	-	0	0	0	0	-
1	1	-	-	-	-	-	-	-	0	-	-	-	-	-

Tabelle 12: *Teilnahmen der Probanden mit ID ≥ 197.*

Motivation	Schulung	Anti Virus	Defacing	Selbstlöschende Datei	Updater	Bounce	Java Update	Login-Fenster	Word Macro	IT Ticket	Nur Link	IT Department	Versichertenkarte	Anti Virus
0	0	-	-	-	-	-	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	3	-	-	-	-	-	-	-	-
0	0	-	-	-	-	-	1	-	1	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	1	-	0	-	-	-	-	1
0	0	-	-	-	-	0	1	-	-	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	1	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	1	1	1	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	0
0	0	-	-	-	-	0	-	1	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	1	1	0	-	-	-	-	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	1	0	-	-	-	-	2
0	0	-	-	-	-	-	1	1	0	-	-	-	-	2

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 12.

0	0	-	-	-	-	-	1	1	-	-	-	-	-	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	1	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	1	1	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	1	-	0	-	-	-	-	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	-	-	-	-	-	0
0	0	-	-	-	-	-	-	-	-	-	-	-	-	0
0	0	-	-	-	-	-	1	-	0	-	-	-	-	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	1	0	-	-	-	-	1
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	0	-	-	-	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	1	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	-	-	0	-	-	-	-	3
0	0	-	-	-	-	0	-	1	0	-	-	-	-	-

Fortsetzung auf der nächsten Seite.

Fortsetzung von Tabelle 12.

0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	3	-	-	0	-	-	-	-	3
0	0	-	-	-	-	-	-	-	-	-	-	-	-	0
0	0	-	-	-	-	0	-	-	0	-	-	-	-	0
0	0	-	-	-	-	-	1	-	0	-	-	-	-	3
0	0	-	-	-	-	-	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	0	-	-	0	-	-	-	-	-
0	0	-	-	-	-	-	-	-	0	-	-	-	-	-

Literaturverzeichnis

[Abr16] Abrams, Lawrence: *New Locky version adds the .Zepto Extension to Encrypted Files*. Bleeping Computer LLC. 2016. URL: https://www.bleepingcomputer.com/news/security/new-locky-version-adds-the-zepto-extension-to-encrypted-files (besucht am 08. 12. 2020).

[AF13] Akhawe, Devdatta & Felt, Adrienne P.: „Alice in Warningland: A Large-Scale Field Study of Browser Security Warning Effectiveness". In: *Presented as Part of the 22nd USENIX Security Symposium (USENIX Security 13)*. 2013.

[Als+12] Alseadoon, Ibrahim; Chan, Taizan; Foo, Ernest & Nieto, Juan G.: „Who Is More Susceptible to Phishing Emails?: A Saudi Arabian Study". In: *Proceedings of the Australasian Conference on Information Systems (ACIS)*. 2012.

[Ana+07] Anandpara, Vivek; Dingman, Andrew; Jakobsson, Markus; Liu, Debin & Roinestad, Heather: „Phishing IQ Tests Measure Fear, Not Ability". In: *Financial Cryptography and Data Security*. Hrsg. von Sven Dietrich & Rachna Dhamija. Springer Berlin Heidelberg, 2007.

[Art09] Arthur, Charles: *Facebook hit by phishing attack*. The Guardian. 2009. URL: https://www.theguardian.com/technology/2009/apr/30/facebook-phishing-scam (besucht am 07. 12. 2020).

[BDS19] BDSG: *Bundesdatenschutzgesetz vom 30. Juni 2017 (BGBl. I S. 2097), das durch Artikel 12 des Gesetzes vom 20. November 2019 (BGBl. I S. 1626) geändert worden ist*. 2019. URL: https://www.gesetze-im-internet.de/bdsg_2018.

[BEd09] Ben-Kiki, Oren; Evans, Clark & döt Net, Ingy: *YAML Ain't Markup Language (YAML^{TM}) Version 1.2*. 2009. URL: https://yaml.org/spec/1.2/spec.html (besucht am 11. 01. 2020).

[Bet20] BetrVG: *Betriebsverfassungsgesetz in der Fassung der Bekanntmachung vom 25. September 2001 (BGBl. I S. 2518), das zuletzt durch Artikel 6 des Gesetzes vom 20. Mai 2020 (BGBl. I S. 1044) geändert worden ist*. Bundesgesetz. 2020. URL: https://www.gesetze-im-internet.de/betrvg.

[BG15] Ben-Asher, Noam & Gonzalez, Cleotilde: „Effects of Cyber Security Knowledge on Attack Detection". In: *Computers in Human Behavior*, 2015.

[BGB20] BGB: *Bürgerliches Gesetzbuch in der Fassung der Bekanntmachung vom 2. Januar 2002 (BGBl. I S. 42, 2909; 2003 I S.738), das zuletzt durch Artikel 2 des Gesetzes vom 12. November 2020 (BGBl. I S. 2392) geändert worden ist*. Bundesgesetz. 2020. URL: https://www.gesetze-im-internet.de/bgb.

[BGL17] Benenson, Zinaida; Gassmann, Freya & Landwirth, Robert: „Unpacking Spear Phishing Susceptibility". In: *Proceedings of the Financial Cryptography and Data Security*. 2017.

[Bie+18] Bieker, Felix; Mohammadi, Linda; Zwingelberg, Harald; Hey, Tim & Ortner, Robert: *Handlungsempfehlungen. Projekt ITS.APT – Dokument 2.4*. Projektbericht. Westfälische Wilhelms-Universität Münster, Institut für Informations- Telekommunikations- und Medienrecht (ITM) und Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD), 2018. URL: https://www.datenschutzzentrum.de/uploads/projekte/its-apt/ITS.APT_D2.4_Handlungsempfehlungen.pdf.

[BPJ13] Boynton, Marcella H.; Portnoy, David B. & Johnson, Blair T.: „Exploring the Ethics and Psychological Impact of Deception in Psychological Research". In: *IRB: Ethics & Human Research*, 2013.

[Bra+06] Bratko, Andrej; Cormack, Gordon V; Filipič, Bogdan; Lynam, Thomas R & Zupan, Blaž: „Spam Filtering using Statistical Data Compression Models". In: *Journal of Machine Learning Research*, 2006.

[Bra14] Brandom, Russell: *TweetDeck vulnerability lets attackers execute code remotely*. The Verge – Vox Media LLC. 2014. URL: https://www.theverge.com/2014/6/11/5800370/tweetdeck-vulnerability-lets-attackers-execute-code-remotely (besucht am 31. 10. 2020).

[Bre92] Brehmer, Berndt: „Dynamic Decision Making: Human Control of Complex Systems". In: Acta Psychologica, 1992.

[Buc+11] Buchanan, Elizabeth; Aycock, John; Dexter, Scott; Dittrich, David & Hvizdak, Erin: „Computer Science Security Research and Human Subjects: Emerging Considerations for Research Ethics Boards". In: *Journal of Empirical Research on Human Research Ethics*, 2011.

[Bun18] Bundesamt für Sicherheit in der Informationstechnik (BSI): *Zuordnungstabelle ISO zum modernisierten IT-Grundschutz*. 2018. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/Zuordnung_ISO_und_modernisierter_IT_Grundschutz.html.

[Bun19] Bundesamt für Sicherheit in der Informationstechnik (BSI): *IT-Grundschutz – Umsetzungshinweise*. 2019.

[Bun20a] Bundesamt für Bevölkerungsschutz und Katastrophenhilfe (BBK): *Kritische Infrastrukturen*. 2020. URL: https://www.kritis.bund.de/DE/AufgabenundAusstattung/KritischeInfrastrukturen/kritischeinfrastrukturen_node.html (besucht am 28. 10. 2020).

[Bun20b] Bundesamt für Sicherheit in der Informationstechnik (BSI): *IT-Grundschutz – Umsetzungshinweise*. 2020. URL: https://www.bsi.bund.de/DE/Themen/ITGrundschutz/ITGrundschutzKompendium/Umsetzungshinweise/Umsetzungshinweise_node.html (besucht am 21. 10. 2020).

[Bun20c] Bundesamt für Sicherheit in der Informationstechnik (BSI): *IT-Grundschutz-Kompendium*. 2020. URL: https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Grundschutz/Kompendium/IT_Grundschutz_Kompendium_Edition2020.pdf.

[Bun20d] Bundesamt für Sicherheit in der Informationstechnik (BSI): *Wie gefährlich ist Phishing?* BSI. 2020. URL: https://www.bsi-fuer-buerger.de/BSIFB/DE/Risiken/SpamPhishingCo/Phishing/Gefahr_

von_Phishing/welche-gefahr-geht-von-phishing-aus_node.html (besucht am 28. 01. 2020).

[Bus+16] Busch, Marc; Fröhlich, Peter; Van der Sype, Yung S.; Hochleitner, Christina; Reisinger, Michaela & Tscheligi, Manfred: „Ethical Implications and Consequences of Phishing Studies in Organizations – An Empirical Perspective". In: *Proceedings of the 26th Conference on Human Factors in Computing Systems (CHI)*. ACM, 2016.

[Cap+14] Caputo, Deanna D.; Pfleeger, Shari L.; Freeman, Jesse D. & Johnson, Eric M.: „Going Spear Phishing: Exploring Embedded Training and Awareness". In: *IEEE Symposium on Security and Privacy*, 2014.

[CI17] Cortes, Juan & Idrizovic, Esmid: *FreeMilk: A Highly Targeted Spear Phishing Campaign*. Palo Alto Networks Inc. 2017. URL: https://unit42.paloaltonetworks.com/unit42-freemilk-highly-targeted-spear-phishing-campaign (besucht am 07. 12. 2020).

[Cla07] Clayton, Richard: „Insecure Real-World Authentication Protocols (or Why Phishing is so Profitable)". In: *Security Protocols*. Springer Berlin Heidelberg, 2007.

[CLC17] Choi, Kwan; Lee, Ju-lak & Chun, Yong-tae: „Voice Phishing Fraud and Its Modus Operandi". In: *Security Journal*, 2017. Hrsg. von Bonnie S. Fisher & Martin Gill.

[Com16] Community, Ruby: *Ruby 2.4.0 – Regexp – Capturing*. 2016. URL: https://docs.ruby-lang.org/en/2.4.0/regexp_rdoc.html#label-Capturing (besucht am 13. 01. 2020).

[Cro+13] Crossler, Robert E.; Johnston, Allen C.; Lowry, Paul B.; Hu, Qing; Warkentin, Merrill & Baskerville, Richard: „Future Directions for Behavioral Information Security Research". In: *Computers & Security*, 2013.

[CWK05] Chan, Mark; Woon, Irene & Kankanhalli, Atreyi: „Perceptions of Information Security at the Workplace: Linking Information Security Climate to Compliant Behavior". In: Journal of Information Privacy and Security, 2005.

[DB16] Döring, Nicola & Bortz, Jürgen: *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Unter Mitarb. von Sandra Pöschl-Günther. Springer-Lehrbuch. Springer, 2016.

[DCF07] Dodge, Ronald C.; Carver, Curtis & Ferguson, Aaron J.: „Phishing for User Security Awareness". In: *Computers & Security*, 2007.

[DHC06] Downs, Julie S.; Holbrook, Mandy B. & Cranor, Lorrie Faith: „Decision Strategies and Susceptibility to Phishing". In: *Proceedings of the Second Symposium on Usable Privacy and Security - SOUPS '06*. ACM Press, 2006.

[DKK14] Dewan, Prateek; Kashyap, Anand & Kumaraguru, Ponnurangam: „Analyzing Social and Stylometric Features to Identify Spear Phishing Emails". In: *Proceedings of the 9th APWG Symposium on Electronic Crime Research (eCrime)*. 2014. arXiv: 1406.3692.

[Eck13] Eckert, Claudia: *IT-Sicherheit. Konzepte – Verfahren – Protokolle*. Oldenbourg Wissenschaftsverlag GmbH, 2013.

[Edw58] Edwards, Allen L.: „The Social Desirability Variable in Personality Assessment and Research". In: *Academic Medicine*, 1958.

[End88] Endsley, Mica R.: „Situation Awareness Global Assessment Technique (SAGAT)". In: *Proceedings of the IEEE 1988 National Aerospace and Electronics Conference*. IEEE, 1988.

[EUE09] Eminağaoğlu, Mete; Uçar, Erdem & Eren, Şaban: „The Positive Outcomes of Information Security Awareness Training in Companies – A Case Study". In: *Information Security Technical Report*, 2009.

[FG06] Ford, Richard & Gordon, Sarah: „Cent, Five Cent, Ten Cent, Dollar: Hitting Botnets Where It Really Hurts". In: *Proceedings of the 2006 Workshop on New Security Paradigms (NSPW)*. 2006.

[FH07] Florencio, Dinei & Herley, Cormac: „A Large-Scale Study of Web Password Habits". In: *Proceedings of the 16th International Conference on World Wide Web*. ACM Press, 2007.

[FJ07] Finn, Peter & Jakobsson, Markus: „Designing and Conducting Phishing Experiments". In: *IEEE Technology and Society Magazine, Special Issue on Usability and Security*, 2007.

[Gai08] Gaillard, Anthony W.K.: „Concentration, Stress and Performance“. In: *Performance Under Stress*. Hrsg. von James L. Szalma & Peter A. A. Hancock. Aldershot: Ashgate, 2008.

[Gar+07] Garera, Sujata; Provos, Niels; Chew, Monica & Rubin, Aviel D.: „A Framework for Detection and Measurement of Phishing Attacks“. In: *Proceedings of the 2007 ACM Workshop on Recurring Malcode (WORM)*. ACM, 2007.

[GDP06] GDPR: *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing — Directive 95/46/EC (General Data Protection Regulation)*. EU Verordnung. 2006. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[GK11] Gellens, R. & Klensin, J.: *Message Submission for Mail*. RFC 6409 (INTERNET STANDARD). Request for Comments. Internet Engineering Task Force (IETF), 2011. URL: http://www.ietf.org/rfc/rfc6409.txt.

[Gre+18] Greene, Kristen; Steves, Michelle; Theofanos, Mary & Kostick, Jennifer: „User Context: An Explanatory Variable in Phishing Susceptibility“. In: *Proceedings 2018 Workshop on Usable Security*. Internet Society, 2018.

[Gré15] Gréaux, Scott: *PhishMe: Disrupting Cyber Attack Detection & Response*. BlackHat. 2015. URL: https://www.blackhat.com/docs/webcast/01222015-phishme-disrupting-cyber-attack.pdf.

[HAO11] Hagen, Janne; Albrechtsen, Eirik & Ole Johnsen, Stig: „The Long-term Effects of Information Security E-learning on Organizational Learning“. In: *Information Management & Computer Security*, 2011.

[Has+05] Hasle, Hågen; Kristiansen, Yngve; Kintel, Ketil & Snekkenes, Einar: „Measuring Resistance to Social Engineering“. In: *Information Security Practice and Experience*. Hrsg. von Robert H. Deng; Feng Bao; HweeHwa Pang & Jianying Zhou. Bearb. von David Hutchison u. a. Springer Berlin Heidelberg, 2005.

[Has05] Hassan, Eman: „Recall Bias Can Be a Threat to Retrospective and Prospective Research Designs“. In: *The Internet Journal of Epidemiology*, 2005.

[Häu15] Häussinger, Felix: „Studies on Employees' Information Security Awareness“. Diss. Niedersächsische Staats-und Universitätsbibliothek Göttingen, 2015.

[HB14] Hänsch, Norman & Benenson, Zinaida: „Specifying IT Security Awareness“. In: *2014 25th International Workshop on Database and Expert Systems Applications*. IEEE, 2014.

[Hel13] Heller, Klaus: *Bekanntmachung des Bundesministeriums für Bildung und Forschung von Richtlinien zur Förderung von Forschungsinitiativen auf dem Gebiet der „IT-Sicherheit für Kritische Infrastrukturen“ im Rahmen des Förderprogramms „IKT 2020 – Forschung für Innovationen“*. Bundesministeriums für Bildung und Forschung. 2013. URL: https://www.bmbf.de/foerderungen/bekanntmachung-878.html (besucht am 18. 12. 2020).

[Hen09] Hensher, David A.: *Hypothetical Bias, Choice Experiments and Willingness to Pay*. Working Paper. University of Sydney – Institute of Transport and Logistics Studies, 2009.

[Her12] Herley, Cormac: „Why Do Nigerian Scammers Say They are From Nigeria?“ In: *WEIS*, 2012.

[Hey+16] Hey, Tim; Ortner, Robert; Jensen, Meiko & Obersteller, Hannah: *Arbeitsrechtliche Risikoabschätzung. Projekt ITS.APT Delivarable 2.1*. Projektbericht. Westfälische Wilhelms-Universität Münster, Institut für Informations- Telekommunikations- und Medienrecht (ITM) und Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD), 2016.

[HMN15] Halevi, Tzipora; Memon, Nasir & Nov, Oded: „Spear-Phishing in the Wild: A Real-World Study of Personality, Phishing Self-Efficacy and Vulnerability to Spear-Phishing Attacks“. In: *SSRN Electronic Journal*, 2015.

[HO16] Hey, Tim & Ortner, Robert: *Haftungsrechtliche Risikoabschätzung. Projekt ITS.APT Delivarable 2.2*. Projektbericht. Westfälische Wilhelms-Universität Münster, Institut für Informations-Telekommunikations- und Medienrecht, 2016.

[Hua+14] Huang, Danny Yuxing u. a.: „Botcoin: Monetizing Stolen Cycles". In: *Proceedings of the Network and Distributed System Security Symposium (NDSS)*. 2014.

[Int13] International Organization for Standardization: *ISO/IEC 27001:2013 Information Technology — Security Techniques — Information Security Management Systems — Requirements*. 2013.

[Int14] Intelligence, Microsoft Security: *Win32/Cribit*. Microsoft. 2014. url: https://www.microsoft.com/en-us/wdsi/threats/malware-encyclopedia-description?Name=Win32/Cribit#symptoms (besucht am 08. 12. 2020).

[Jag+07] Jagatic, Tom N.; Johnson, Nathaniel A.; Jakobsson, Markus & Menczer, Filippo: „Social Phishing". In: *Communications of the ACM*, 2007.

[Jak05] Jakobsson, Markus: „Modeling and Preventing Phishing Attacks". In: *Financial Cryptography*. 2005.

[Jak07] Jakobsson, Markus: „The Human Factor in Phishing". In: *Privacy & Security of Consumer Information*, 2007.

[JF08] Jakobsson, Markus & Finn, Peter: „Why and How to Perform Fraud Experiments". In: *IEEE Security & Privacy*. 2008.

[JM06] Jakobsson, Markus & Myers, Steven: *Phishing and Countermeasures: Understanding the Increasing Problem of Electronic Identity Theft*. Wiley, 2006.

[JO16] Jensen, Meiko & Obersteller, Hannah: *Datenschutzrechtliche Betrachtung. Projekt ITS.APT – Dokument 2.3*. Projektbericht. Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein (ULD), 2016.

[JR06] Jakobsson, Markus & Ratkiewicz, Jacob: „Designing Ethical Phishing Experiments: A Study of (ROT13) rOnl Query Features“. In: *Proceedings of the 15th International Conference on World Wide Web*. 2006.

[Kab94] Kabay, Michel: „Psychosocial Factors in the Implementation of Information Security Policy“. In: *The EDP Audit, Control, and Security Newsletter (EDPACS)*, 1994.

[KD12] Kenneally, Erin & Dittrich, David: *The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research*. Commission Rerport. U.S. Department of Homeland Security, 2012.

[Kel+12] Kelley, Christopher M.; Hong, Kyung Wha; Mayhorn, Christopher B. & Murphy-Hill, Emerson: „Something Smells Phishy: Exploring Definitions, Consequences, and Reactions to Phishing“. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2012.

[Kha+11] Khan, Bilal; Alghathbar, Khaled S.; Nabi, Seyed I. & Kahn, Muhammad K.: „Effectiveness of Information Security Awareness Methods Based on Psychological Theories“. In: *African Journal of Business Management*, 2011.

[KK05] Kirda, Engin & Kruegel, Christopher: „Protecting Users Against Phishing Attacks with AntiPhish“. In: *Proceedings of the 29th Annual International Computer Software and Applications Conference (COMPSAC)*. IEEE Computer Society, 2005.

[KK06] Kruger, Hennie A. & Kearney, Wayne D.: „A Prototype for Assessing Information Security Awareness“. In: *Computers & Security*, 2006.

[Kum+07] Kumaraguru, Ponnurangam; Rhee, Yong; Acquisti, Alessandro; Cranor, Lorrie F.; Hong, Jason & Nunge, Elizabeth: „Protecting People from Phishing: The Design and Evaluation of an Embedded Training Email System“. In: *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI)*. 2007.

[Kum+08] Kumaraguru, Ponnurangam; Sheng, Steve; Acquisti, Alessandro; Cranor, Lorrie F. & Hong, Jason: „Lessons from a Real World Evaluation of Anti-Phishing Training“. In: *2008 eCrime Researchers Summit*. IEEE, 2008.

[Lab68] Labovitz, Sanford: „Criteria for Selecting a Significance Level: A Note on the Sacredness of. 05“. In: *The American Sociologist*, 1968.

[Lav10] Lavie, Nilli: „Attention, Distraction, and Cognitive Control Under Load“. In: *Current Directions in Psychological Science*, 2010.

[Leb+13] Lebek, Benedikt; Uffen, Jorg; Breitner, Michael H.; Neumann, Markus & Hohler, Bernd: „Employees' Information Security Awareness and Behavior: A Literature Review“. In: *2013 46th Hawaii International Conference on System Sciences*. IEEE, 2013.

[Leb+14] Lebek, Benedikt; Uffen, Jörg; Neumann, Markus; Hohler, Bernd & Breitner, Michael H.: „Information Security Awareness and Behavior: A Theory-Based Literature Review“. In: *Management Research Review*, 2014.

[LFC12] Lai, Jerry; Fidler, Fiona & Cumming, Geoff: „Subjective p Intervals: Researchers Underestimate the Variability of p Values Over Replication“. In: *Methodology*, 2012.

[McC+16] McCormac, Agata; Calic, Dragana; Parsons, Kathryn; Zwaans, Tara; Butavicius, Marcus & Pattison, Malcolm: „Test-Retest Reliability and Internal Consistency of the Human Aspects of Information Security Questionnaire (HAIS-Q)“. In: 2016.

[McL95] McLean, Kevin: „Control Concepts — Who's Buying?“ In: *Computer Audit Update*, 1995.

[MdBW12] McCambridge, Jim; de Bruin, Marijn & Witton, John: „The Effects of Demand Characteristics on Research Participant Behaviours in Non-Laboratory Settings: A Systematic Review“. In: *PLOS ONE*, 2012.

[Mic18] Microsoft: *Microsoft Security Servicing Criteria for Windows*. 2018. URL: https://www.microsoft.com/en-us/msrc/windows-security-servicing-criteria (besucht am 11. 01. 2020).

[Mil07] Milletary, Jason: *Technical trends in phishing attacks*. Techn. Ber. CERT Coordination Center, 2007. URL: https://us-cert.cisa.gov/sites/default/files/publications/phishing_trends0511.pdf.

[MJ92] McCrae, Robert R. & John, Oliver P.: „An Introduction to the Five-Factor Model and Its Applications". In: *Journal of Personality*, 1992.

[MKK08] Medvet, Eric; Kirda, Engin & Kruegel, Christopher: „Visual-similarity-based Phishing Detection". In: *Proceedings of the 4th International Conference on Security and Privacy in Communication Netowrks (SecureComm)*. ACM, 2008.

[Moz16] Mozilla Corporation: *Firefox*. 2016. URL: https://www.mozilla.org/de (besucht am 27. 08. 2020).

[Mül17] Müller, Silke: *Schulungskonzept – IT-Sicherheit. IT-Security Awareness Penetration Testing (ITS.APT)*. Projektbericht. Universität Duisburg-Essen, 2017.

[MW47] Mann, Henry B. & Whitney, Donald R.: „On a test of whether one of two random variables is stochastically larger than the other". In: *The Annals of Mathematical Atatistics*, 1947.

[MWE14] McCambridge, Jim; Witton, John & Elbourne, Diana R.: „Systematic Review of the Hawthorne Effect: New Concepts Are Needed to Study Research Participation Effects". In: *Journal of Clinical Epidemiology*, 2014.

[NL14] Nohl, Karsten & Lell, Jakob: *BadUSB – On Accessories that Turn Evil*. BlackHat USA. 2014. URL: https://www.blackhat.com/us-14/briefings.html#badusb-on-accessories-that-turn-evil (besucht am 20. 01. 2021).

[Ori16] Oriyano, Sean-Philip: *CEH^{TM}v9 Certified Ethical Hacker Version 9: Study Guide*. Wiley, 2016.

[Par+13a] Parsons, Kathryn; McCormac, Agata; Butavicius, Marcus; Pattinson, Malcolm & Jerram, Cate: „The Development of the Human Aspects of Information Security Questionnaire (HAIS-Q)". In: 2013.

[Par+13b] Parsons, Kathryn; McCormac, Agata; Pattinson, Malcolm; Butavicius, Marcus & Jerram, Cate: „Phishing for the Truth: A Scenario-Based Experiment of Users' Behavioural Response to Emails". In: *Proceedings of the IFIP International Information Security Conference*. Hrsg. von Lech J. Janczewski; Henry B. Wolfe & Sujeet Shenoi. 2013.

[Par+14] Parsons, Kathryn; McCormac, Agata; Butavicius, Marcus; Pattinson, Malcolm & Jerram, Cate: „Determining Employee Awareness Using the Human Aspects of Information Security Questionnaire (HAIS-Q)". In: *Computers & Security*, 2014.

[Par+17] Parsons, Kathryn; Calic, Dragana; Pattinson, Malcolm; Butavicius, Marcus; McCormac, Agata & Zwaans, Tara: „The Human Aspects of Information Security Questionnaire (HAIS-Q): Two Further Validation Studies". In: *Computers & Security*, 2017.

[PL01] Poitevineau, Jacques & Lecoutre, Bruno: „Interpretation of Significance Levels by Psychological Researchers: The .05 Cliff Effect May Be Overstated". In: *Psychonomic Bulletin & Review*, 2001.

[Red16] Reddit Inc.: *TeamViewer has been hacked. They are denying everything and pointing fingers at the users.* 2016. url: https://www.reddit.com/r/technology/comments/4m7ay6/teamviewer_has_been_hacked_they_are_denying/ (besucht am 17. 11. 2020).

[Res78] Research, National Commission for the Protection of Human Subjects of Biomedical and Behavioral: *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Commission Rerport. Department of Health, Education and Welfare (DHEW), 1978.

[RF18] Resnik, David B. & Finn, Peter R.: „Ethics and Phishing Experiments". In: *Science and Engineering Ethics*, 2018.

[RvdH17] Romagna, Marco & van den Hout, Niek J.: „Hacktivism and Website Defacement: Motivations, Capabilities and Potential Threats". In: *Proceedings of the Virus Bulletin Conference*. 2017.

[SB14] Stobert, Elizabeth & Biddle, Robert: „The Password Life Cycle: User Behaviour in Managing Passwords". In: *Proceedings of the 10th*

Symposium On Usable Privacy and Security ($\{$SOUPS$\}$ 2014. 2014.

[SCC02] Shadish, William R.; Cook, Thomas D. & Campbell, Donald T.: *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin, 2002. 623 S.

[Sch06] Schulze, Tillmann: *Bedingt abwehrbereit. Schutz kritischer Informations-Infrastrukturen in Deutschland und den USA*. VS Verlag für Sozialwissenschaften, 2006.

[Sch10] Schonfeld, Erick: *What Happened To bit.ly's Market Share?* 2010. URL: https://techcrunch.com/2010/01/06/bit-ly-market-share-2/ (besucht am 19. 10. 2020).

[Sco15] Scott, David W.: *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons, 2015.

[SGT19] Steves, Michelle P.; Greene, Kristen K. & Theofanos, Mary F.: „A Phish Scale: Rating Human Phishing Message Detection Difficulty". In: *Proceedings 2019 Workshop on Usable Security*. Internet Society, 2019.

[SH99] Srisuresh, P. & Holdrege, M.: *IP Network Address Translator (NAT) Terminology and Considerations*. RFC 2663 (Informational). Request for Comments. Internet Engineering Task Force (IETF), 1999. URL: http://www.ietf.org/rfc/rfc2663.txt.

[Sha05] Shafranovich, Y.: *Common Format and MIME Type for Comma-Separated Values (CSV) Files*. RFC 4180 (Informational). Request for Comments. Updated by RFC 7111. Internet Engineering Task Force (IETF), 2005. URL: http://www.ietf.org/rfc/rfc4180.txt.

[Sin06] Singh, Amit: *Mac OS X Internals. A Systems Approach*. Addison Wesley Professional, 2006.

[SKD07] Steyen, Tjaart; Kruger, Hennie A. & Drevin, Lynette: „Identity Theft — Empirical Evidence from a Phishing Exercise". In: *Proceedings of the International Information Security Conference (IFIP)*. Springer, 2007.

[Smi14] Smith OficialKLS, Kev: *wtf?!* Twitter Inc. 2014. URL: https://twitter.com/OfficialKLS/status/476755204868935682 (besucht am 31. 10. 2020).

[SSS18] Schofield, McLean; Sharkey, Kent & Satran, Michael: *Installation Package*. Microsoft. 2018. URL: https://docs.microsoft.com/en-us/windows/win32/msi/installation-package (besucht am 08. 01. 2021).

[Sta+05] Stanton, Jeffrey M.; Stam, Kathryn R.; Mastrangelo, Paul & Jolton, Jeffrey: „Analysis of End User Security Behaviors". In: *Computers & Security*, 2005.

[Sta05] Standardization, International Organization for: *ISO/IEC27002 Information Technology — Security Techniques — Code of Practice for Information Security Management*. 2005.

[Sun+09] Sunshine, Joshua; Egelman, Serge; Almuhimedi, Hazim; Atri, Neha & Cranor, Lorrie F.: „Crying Wolf: An Empirical Study of SSL Warning Effectiveness." In: *USENIX Security Symposium*. 2009.

[SW21] Sykosch, Arnold & Wübbeling, Matthias: *IT-Security Awareness Penetration Environment*. GitLab. 2021. URL: https://gitlab.com/itsape (besucht am 13. 01. 2020).

[Swe02] Sweeney, Latanya: „K-Anonymety: A Model for Protecting Privacy". In: *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2002.

[Syk+20] Sykosch, Arnold; Doll, Christian; Wübbeling, Matthias & Meier, Michael: „Generalizing the Phishing Principle: Analyzing User Behavior in Response to Controlled Stimuli for IT Security Awareness Assessment". In: *Proceedings of the 15th International Conference on Availability, Reliability and Security*. ACM, 2020.

[Sza14] Szappanos, Gabor: *VBA is not dead!* Virus Bulletin. 2014. URL: https://www.virusbulletin.com/virusbulletin/2014/07/vba-not-dead (besucht am 13. 11. 2020).

[Tab18] Taber, Keith S.: „The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education". In: *Research in Science Education*, 2018.

[TCF10] Talib, Shuhaili; Clarke, Nathan L. & Furnell, Steven M.: „An Analysis of Information Security Awareness within Home and Work Environments“. In: *2010 International Conference on Availability, Reliability and Security*. IEEE, 2010.

[Tea20] TeamViewer Germany GmbH: *Das ist TeamViewer*. 2020. URL: https://www.teamviewer.com/de/produkte/teamviewer/ (besucht am 17.11.2020).

[Theo7] The European Network and Information Security Agency (ENISA): *Information Security Awareness Initiatives: Current Practice and the Measurement of Success*. Studie. ENISA, 2007.

[Theo8] The European Commission: „On the identification and designation of European critical infrastructures and the assessment of the need to improve their protection — Council Directive 2008/114/EC“. In: *Official Journal of the European Union*, 2008. URL: https://eur-lex.europa.eu/eli/dir/2008/114/oj.

[The10] The European Network and Information Security Agency (ENISA): *The new users' guide: How to raise information security awareness*. 2010.

[TM17] Tröster, Benedikt & Matula, Oliver: *Schulungsprotokoll (D4.6). IT-Security Awareness Penetration Testing (ITS.APT)*. Projektbericht. ERNW Enno Rey Netzwerke GmbH, 2017.

[Tso+08] Tsohou, Aggeliki; Kokolakis, Spyros; Karyda, Maria & Kiountouzis, Evangelos: „Investigating Information Security Awareness: Research and Practice Gaps“. In: *Information Security Journal: A Global Perspective*, 2008.

[Una16] Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein: *Datenschutzrecht in Schleswig-Holstein*. Gesetzessammlung. (ULD), 2016. URL: https://www.datenschutzzentrum.de/uploads/gesetze/Band_1-Datenschutzrecht-2016-6.Auflage-web.pdf.

[VER20] VERIZON: *2020 Data Breach Investigations Report*. Techn. Ber. 2020.

[Ves11] Veseli, Ilirjana: „Measuring the Effectiveness of Information Security Awareness Program“. Magisterarb. Gjøvik University College: Department of Computer Science and Media Technology, 2011.

[VvS04] Vroom, Cheryl & von Solms, Rossouw: „Towards Information Security Behavioural Compliance“. In: *Computers & Security*, 2004.

[W3C18] W3C Working Group: *CSS Snapshot 2018 – Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification*. 2018. URL: https://www.w3.org/TR/CSS2/colors.html#propdef-background-image (besucht am 20. 10. 2020).

[WBS08] Workman, Michael; Bommer, William H. & Straub, Detmar: „Security Lapses and the Omission of Information Security Measures: A Threat Control Model and Empirical Test“. In: *Computers in Human Behavior*, 2008.

[WH03] Wilson, Mark & Hash, Joan: *Building an Information Technology Security Awareness and Training Program*. NIST Special Publication. National Institute of Standards and Technology, 2003. URL: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-50.pdf (besucht am 06. 01. 2020).

[WHP11] Wolf, Michael; Haworth, Dwight & Pietron, Leah: „Measuring An Information Security Awareness Program“. In: *Review of Business Information Systems (RBIS)*, 2011.

[Wri+10] Wright, Ryan; Chakraborty, Suranjan; Basoglu, Asli & Marett, Kent: „Where Did They Go Right? Understanding the Deception in Phishing Communications“. In: *Group Decision and Negotiation*, 2010.

[Xia14] Xiao, C.: *WIRELURKER: A New Era in iOS and OS x Malware*. Threat report. PALO ALTO NETWORKS: unit42, 2014.

[Zor10] Zorz, Zeljka: *World of Warcraft phishing scams*. Help Net Security. 2010. URL: https://www.helpnetsecurity.com/2010/09/29/world-of-warcraft-phishing-scams (besucht am 07. 12. 2020).

Abbildungsverzeichnis

Tabellenverzeichnis