Original citation:

Permanent WRAP URL:
http://wrap.warwick.ac.uk/94777

Copyright and reuse:
The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions. Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:
© 2017 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting /republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

A note on versions:
The version presented here may differ from the published version or, version of record, if you wish to cite this item you are advised to consult the publisher’s version. Please see the ‘permanent WRAP url’ above for details on accessing the published version and note that access may require a subscription.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Harjinder Singh Lallie*, Kurt Debattista†, Jay Bal‡
*Cyber Security Centre, WMG, University of Warwick, UK
†WMG, University of Warwick, UK
‡Cyber Security Centre, WMG, University of Warwick, UK

Abstract—Perceiving and understanding cyber-attacks can be a difficult task. This problem is widely recognised and well documented, and more effective techniques are needed to aid cyber-attack perception. Attack modelling techniques (AMTs) - such as attack graphs and fault trees, are useful visual aids that can aid cyber-attack perception; however, there is little empirical or comparative research which evaluates the effectiveness of these methods. This paper reports the results of an empirical evaluation between an adapted attack graph method and the fault tree standard to determine which of the two methods is more effective in aiding cyber-attack perception. An empirical evaluation ($n=63$) was conducted through a $3 \times 2 \times 2$ factorial design. Participants from computer-science and non-computer-science backgrounds were divided into an adapted attack graph and fault tree group and then asked to complete three tests which tested the ability to recall, comprehend and apply the attack modelling technique. A mean assessment score (mas) was calculated for each test.

The results show that the adapted attack graph method is more effective at aiding cyber-attack perception when compared with the fault tree method ($p<0.01$). Participants that have a computer science background outperformed other participants when using both methods ($p<0.05$). These results indicate that the adapted attack graph method can be an effective tool for aiding cyber-attack perception amongst experts. The study underlines the need for further comparisons in a broader range of settings involving additional techniques, and suggests several suggestions for further work.

Index Terms—Cyber-attack, attack modelling, cyber-visualisation, attack graph, fault tree, attack tree

I. INTRODUCTION

Recent well known attacks such as the WannaCry ransomware attack [1] have highlighted problems relating to the ability of decision makers to perceive cyber-security - in particular cyber-attacks [2]. The problem of perception in the present study relates to the ability to assess and understand the sequence of events that led to a cyber-attack. This problem applies to expert computing/IT practitioners - who have to make real time incident response decisions as well as non-experts, such as CEOs - who are part of the decision-making process and may not fully understand the technical implications. CEOs have low cyber-literacy levels - with 91% reporting problems with interpreting cyber-security reports [3]. Decision makers can be poorly advised and under-prepared to tackle cyber-security challenges [4] and don’t necessarily possess the key knowledge and understanding to drive action [5, 2]. Consequently, there is a view that decision makers consider the cyber-security domain to be perceptually inaccessible [6].

Better techniques are required to aid the understanding of cyber-security among such audiences so as to enable a more effective understanding of risk [7] and better decision making. Attack modelling techniques (AMT) are a method of modelling and visualising the sequence of events that enable a successful cyber-attack on a host or network. AMTs allow analysts to understand the underlying susceptibility of a host or network to a cyber-attack - and in doing so, to identify weaknesses and vulnerabilities in a host or network. Furthermore, these techniques permit decision-makers and other non-experts to form an understanding of potential underlying vulnerabilities and means of preventing them.

Although AMTs serve a useful academic and practical purpose, there has been little or no research on the measurable cognitive impact of AMTs - on both experts and non-experts. This situation exists for many other visual information flow models as well. For example, there exist few if any controlled evaluations comparing data flow charts or status diagrams with other methods of expressing information flow. Quite often there is a presumption that a particular visual method is the ‘best’ mechanism for expressing information flow. This is an important research area, and this paper contributes towards addressing this research gap by presenting the results of an empirical evaluation of attack graphs and fault trees as methods of aiding cyber-attack perception.

Although numerous research papers have deployed attack graphs to represent cyber-attacks, there is no standardised attack graph method. More than fifty self-nominated methods have been adopted in the academic literature to display an attack graph - each of which represents the key aspects of a cyber attack in a subtly different way. The domain space demands the proposal of a validly formulated attack graph method which fully represents the fundamental cyber-attack constructs. Fault trees on the other hand are defined by an international standard [8] and have been used to describe cyber-attacks.

This study compares an adapted attack graph ($aug$) with the fault tree method to determine which of the two methods is more effective at aiding cyber-attack perception. While the two methods bear some conceptual
similarities, they differ in terms of the symbol construction and data flow, and the study attempts to outline which of the two sets of visual structures is more effective at aiding cyber-attack perception. The term ‘effectiveness’ in the present context refers to the ability of a participant to respond correctly to a question requiring the interpretation of the visual syntax of a given AMT. The study finds that the *aag* method is more effective than the fault tree method (*p* < 0.01) at aiding cyber-attack perception. Furthermore, participants with some computer science knowledge demonstrated a higher ranking when using *aag* in comparison with those using the fault tree method (*p* < 0.05) signifying that the *aag* method can be effective in increasing cyber-attack perception amongst experts.

The novelty and contributions of the work presented herein are as follows, the research:

1) Demonstrates that the *aag* can be used as a more effective method of aiding cyber-attack perception amongst expert audiences when compared with fault trees
2) Proposes a methodology that enables researchers to measure the effectiveness of visual information flow methods
3) Outlines initial efforts towards defining a standardised attack graph method.

The rest of this paper is organised as follows. Section II outlines related work. Section III introduces the *aag* method and fault tree methods. Section IV outlines the experimental framework and outlines the methods used, the demographics of the participants and the procedure followed by the participants. Section V provides an analytical overview of the results - particularly highlighting the tests applied to the results to highlight the statistical significance. Section VI explains the results and outlines areas for further research.

II. BACKGROUND

The ability to present security problems in a manner that enables decision makers to effectively perceive the problem is a key security challenge [9]. The visualisation of complex knowledge and information structures helps learners to better perceive complex concepts. Visual methods such as AMTs have considerable value in aiding cyber-attack perception [10, 11]. Fithen et al., [12] outline the benefits of attack visualisation by arguing that AMTs enable non-experts to better understand and interpret attack models with little reference to logical models. Roschke et al., [13] propose that AMTs remove the intellectual burden from security experts who have to understand and evaluate numerous potential options. Effective visual representations enable security personnel to achieve a quick understanding of the problem domain.

A lot of the research into cyber-attack perception appears to have focussed on aiding the cyber situational awareness of experts [14, 15, 16]. However, there is a recognition - expressed largely in the business press, that there is a wide-spread problem of cyber-attack perception amongst non-experts [5, 6] and a recognition that AMTs can aid the assessment and understanding of cyber-attacks. However, there is a dearth of academic research into the benefits of AMTs on cyber-attack perception.

Opdahl and Sindre [17] performed a qualitative evaluation which compared the attack tree method with *misuse cases* in aiding practitioner perception in threat identification. Opdahl and Sindre set out to evaluate the effectiveness of the two techniques - measured as the number of threats found by participants in two sample scenarios, the number/types of threats found by participants; and participant perceptions of the two techniques. The research presented three key findings. The attack tree method was more effective in aiding threat identification when compared with the misuse case method, participant perceptions of two techniques were similar, however, participant perception did not correlate with the performance of the participant in using the selected technique.

The contribution by Flaten and Lund [18] attempted to understand whether attack trees could improve an expert’s understanding of a cyber-security threat. The research harnessed the views of two cyber-security experts who answered a number of questions regarding two attack scenarios - both presented as attack trees. The study found that attack trees are not suited for aiding cyber-attack perception in cyber-security experts. Other than these two studies there appears to have been no other research focussing on the cognitive benefits of AMTs.

The key differences between the approaches of Opdahl and Sindre, and Flaten and Lund and the approach presented herein are as follows. The contribution by Flaten and Lund involved a small number of participants (*n*=2) and considered one AMT. The present study adopts a quantitative approach with 63 participants. The two methods compared by Opdahl and Sindre, and Flaten and Lund (attack trees and misuse cases) are considerably different in terms of their visual syntax. Such a comparison requires a careful analysis and consideration of the impact that differences in the two techniques could have on participants.

The present method compares two conceptually similar techniques. Such an approach allows for a focussed analysis of the two techniques and enables a refined understanding of the differences between the two techniques.

III. ATTACK MODELLING TECHNIQUES

AMTs represent cyber-attacks by using semantic methods (formal languages) and/or visual syntax [72] in the form of a tree/graph/net. The visual representation of an attack - referred to herein as ‘visual syntax’, utilises symbolic modes of expression to visualise one or more of the three *fundamental cyber-attack constructs* which are: the preconditions/postconditions of a cyber-attack (also referred to as a ‘status’); exploits (also referred to as an ‘event’); and pre-condition logic. Examples of these are given in Figure 1 where a precondition is represented as a box, an exploit as an ellipse and precondition logic by the presence or absence of an arc connecting two edges.

An exploit can be represented as a tuple of the form: *(h_s,h_d,v)* wherein source host *h_s* can exploit a vulnerability *v* which exists on a destination host *h_d* [73]. A precondition is one or more host/system statuses that must exist for an exploit to be successful. The postcondition is the state of the host/system once the exploit has
In this example, any one of two exploits (sshd) to be successful. An example of this is provided in Figure 1 wherein two preconditions must exist for the sshd exploit to be applied: sshd(3,1) and user(3). A disjunctive precondition relationship requires for any one or more of the connected preconditions to be fulfilled for the exploit to be successful. An example of this is provided in Figure 1. In this example, any one of two exploits (sshd_bof(1,2) or sshd_bof(3,2)) must be applied to gain user(2) status on the target machine. It is not common for precondition logic to be represented in a given AMT; however, it is important for an analyst to know the relationship between the conditions. Addressing just one of multiple conjunctive preconditions - which an exploit relies upon, can act as a suitable mitigation strategy.

Numerous AMTs have been proposed in the academic literature. A number of the more popular AMTs are listed in Table I and include: attack trees [19, 20], defence trees [26, 27, 28], privilege graphs [56], exploitation graphs or e-graphs, [57, 59], fault trees [33], Petri Net Models [20, 34, 33], alert correlation graphs [48, 49, 51, 13, 52] and influence diagrams [62, 63, 64]. Of these methods, fault trees and Petri nets are the only attack modelling techniques to be defined in an international standard [8, 74]. Some of these methods - for example fault trees and Petri nets, were not designed to represent cyber-attacks, but have been used in the literature to describe cyber-attacks.

### A. Selection of AMT

The present study compares an adapted attack graph method with the fault tree method. Attack graphs and fault trees were chosen for the empirical comparison for reasons of functionality and because they are widely academically accepted.

**Functionality:** One of the key factors in selecting the two methods to be compared was whether the method is able to represent the fundamental cyber-attack constructs without requiring significant modification. Of all the methods outlined in Table I, attack graphs (particularly those described by Barik and Mazumdar [47] and Ghosh and Ghosh [75]) and fault trees are able to represent these constructs with very little or no modification. Methods such as kill chains [70] and the diamond model [71] are not designed to represent the fundamental cyber-attack constructs. Petri nets were designed to enable experts to understand information flow and control in systems and are particularly useful in describing systems that exhibit concurrency and asynchronous behaviour [76]. They are useful for experts but do not lend themselves to easy perception by non experts. It would be unfair to compare such a system with a method that is much more visually expressive.

**Academic acceptance:** Attack graphs and fault trees are an accepted and popular form of attack representation amongst the academic community and have been applied in multiple wide ranging scenarios. Methods such as kill chains and the diamond model are popular amongst the business community but not necessarily amongst the academic community.

The fault tree design is governed by an international standard [8] which defines numerous symbols. A small subset of these symbols are ideal for use in describing a cyber-attack scenario - making the fault tree method ideal for comparison.

Both methods are described more fully in section III-C1 and III-C2.

### B. Visual Structures

The adapted attack graph (aag) method is based on the attack graph representations by Noel et al., [44], Foo et al., [45], Wang et al., [46], and Barik and Mazumdar [47]. The adapted attack graph method is demonstrated in Figures 1 (right), 2 (right) and 3 (right) and the following key explains the icons used therein.

- Preconditions/postconditions: rectangle
- Exploits: ellipse
- Conjunction: arc connecting preconditions
- Disjunction: Absence of an arc connecting preconditions
- Grey rectangle: overall attack goal

The two models are conceptually similar. However, there are some fundamental differences in their visual representation.
syntax structures which are likely to render differences in cognitive perception. These differences are presented in Table II and can be summarised as follows.

- **Symbol usage** Both methods utilise two symbols to represent exploits and preconditions. aag utilises an ellipse for an exploit and a rectangle for a precondition. For the fault tree method it is the other way around.
- **Representation of precondition logic.** The fault tree method utilises two symbols to represent precondition logic. The aag method utilises the presence or absence of an arc to represent conjunction and disjunction respectively.
- **Symbol count.** Correspondingly, the total number of symbols used in the fault tree method is four, in the aag method it is two.
- **Event flow.** Events flow upwards in a fault tree (or rather conceptual reasoning starts at the top) whereas in the aag method (and in attack graphs in general) they flow downwards.

Some of these simple differences such as the inclusion of a specific symbol to represent an OR condition - thereby increasing the total symbol count, and the direction of information flow are likely to impact the results.

**C. Attack Graphs and Fault Trees**

1) **Attack Graphs:** Attack graphs are possibly the most popular AMT - particularly in the academic literature. An attack graph is a mathematical abstraction of attack paths that might be perpetrated against a given system [77, 78]. The graph comprises of nodes which represent exploits/attacks/events and edges which represent a change of status.

The nodes in the graph can represent a range of elements such as:

- An **exploit** that has been or could be applied to the given node [61, 79, 80]
- An **event** such as an access violation [81] or a remote server exploit [82, 51] which forms the necessary stages in an attack
- A **status** or condition attained by an exploit [83]. Examples of this include preconditions/postconditions.

Quite often, preconditions/postconditions are defined as privileges [55, 56].

Edges in an attack graph can be directed - to represent specific transitions, or undirected - to represent a general connection between two nodes and generally represent the perpetration of an exploit. However, edges can also represent: actions [84], preconditions - where a precondition edge \( e = (a, s) \) exists if \( a \) is a precondition of \( s \) [57], or vulnerabilities - where a vulnerability edge represents a vulnerability that a perpetrator could exploit [85, 78].

2) **Fault Trees:** Both fault trees and attack trees find their origins in decision trees which aid decision making through the representation of three elements: a decision node, edge/branch and leaf. Although decision trees have been applied to a computer/cyber security context [86, 87, 88, 89, 90, 40], decision trees were not designed to aid visual consumption, and it is probably for these reasons that methods such as attack trees and fault trees were developed. Schneier recast decision trees in the form of attack trees [19] and may have been influenced by the fault tree method [40].

Attack trees and fault trees are acyclic directed graphs which outline important events and conditions. Events lead to a goal condition which is referred to as an attack goal in an attack tree [19], and an undesirable condition in the case of a fault tree [91].

The symbolic representation of fault trees was standardised by the IEC in 1990, [8], the European Cooperation for Space Standardization [92] and then by the British Standards Institute [93]. However, although the fault tree structure is standardised, there is no agreed method of representing attack trees, and - like attack graphs, there exist numerous - subtly different versions of attack trees.

Fault trees are the most visually expressive AMT because they utilise a wide range of standard symbols to express elements of an attack. Fault trees are used in a number of industries such as in the aerospace industry [94, 95, 96, 97], radioactive waste disposal [98], the automotive industry [99, 100] and in the analysis of failure in computer systems [101, 102, 103]. Although the fault tree standard is a generic standard (not particularly focussing on cyber-security as a target domain), more recently, fault trees have become a popular means of representing cyber-attack [40, 104, 105].

Example fault trees are given in Figures 1, 2 and 3. These examples demonstrate the use of conjunction and disjunction with specific symbols, preconditions as ovals and exploits as rectangles.

3) **Syntactic Structure:** Attack graphs and fault trees comprise of two fundamental elements represented as graph data structures of the form: \( G(V; E) \) [108] which comprises of nodes and edges: \( e \in E \) which represent relationships between the nodes. An attack graph/fault tree is a tuple \( G = (S, \tau, S_0, S_f, L, E) \) where:

- \( S \) is a finite set of states,
- \( \tau \subseteq S \times S \) is a transition relation
- \( S_0 \subseteq S \) is a set of initial states
- \( S_f \subseteq S \) is a set of success states – for example obtaining root or user privileges on a particular host.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Fault tree</th>
<th>Adapted Attack Graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exploit</td>
<td>rsh(1.2)</td>
<td>rsh(1.2)</td>
</tr>
<tr>
<td>Precondition</td>
<td>user(2)</td>
<td>user(2)</td>
</tr>
<tr>
<td>And</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Or</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Flow</td>
<td>Bottom to top</td>
<td>Top to Bottom</td>
</tr>
</tbody>
</table>

Table II: Visual Syntactic Differences between the Fault Tree and Adapted Attack Graph Methods
<table>
<thead>
<tr>
<th>Test</th>
<th>Lower order cognitive skill</th>
<th>Test Description</th>
<th>Scenario reference</th>
<th>Sample question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Knowledge Recall</td>
<td>Multiple Choice select one answer</td>
<td>scenario 1 [47], 4 questions</td>
<td>\textbf{“What are the necessary exploits for an attacker to be able to achieve user access on host 2”, also see Figure 4}</td>
</tr>
<tr>
<td>2</td>
<td>Comprehension</td>
<td>Select correct scenario from a heatmap</td>
<td>scenario 2 [75] 4 questions</td>
<td>\textbf{“Study the image below and select the exploit(s) which result in the attacker gaining user access status on host 2.”}</td>
</tr>
<tr>
<td>3</td>
<td>Application</td>
<td>Multiple Choice, read scenario and select one from three heat maps</td>
<td>scenario 3 [106, 107] 4 questions</td>
<td>\textbf{Study the figure below and select the figure that most accurately describes the following scenario: “The stunner virus is installed when a new services.exe file and a new s7otidwsdx.dll file are installed. Before these can be installed, the following preconditions must be met. The target has to have the RPC vulnerability, the target has to be running the Step7 application, and the target has to be a Stimulation PLC.”}</td>
</tr>
</tbody>
</table>

Table III: Description of Study Scenarios and Tests

- $L : S \rightarrow 2^{AP}$ is a labelling of states with a set of atomic propositions ($AP$)
- $E$ is a finite set of exploits which connect the transition between two states

IV. EMPIRICAL EVALUATION

A. Design

The empirical evaluation uses three independent variables ($test, AMT, background$) and one dependent variable ($mean assessment score - mas$).

The study aims to evaluate whether any of the three independent variables ($test, AMT$ and $background$) influence cyber-attack perception.

The $test$ is a within participant independent variable which represents the three questions asked to participants. Cyber-attack assessment and understanding was measured using a three phase test which demands the demonstration of increasingly complex cognitive skills. Table III provides an overview of its characteristics. Each test corresponded to one of the three lower levels (knowledge, comprehension and application) of Bloom’s Taxonomy of educational objectives [109]. Bloom’s taxonomy has been used in numerous academic fields to assess both lower and higher order cognitive skills. The result of the test is measured as a mean assessment score $mas$ - the dependent variable.

The wording of the questions was carefully prepared to correspond with guidance provided on how to frame questions and utilise keywords so as to correspond with levels within the taxonomy (for example, see [110]). The test framework - including question samples, are provided in Table III.

The $background$ independent variable is divided into two groups: $cs$ and $oth$. The $cs$ group have a computer science background. These participants have either studied computer science at undergraduate level, or have more than five years of work experience in the computing industry. The $other$ group are all other participants.

There are two $AMT$ groups: $aag$ and $ft$ - each presented with an adapted attack graph or fault tree scenario respectively.

The following two hypotheses were established:

$H1_1$ The selection of $AMT$ influences response to $mas$

$H2_1$ The selection of $background$ influences response to $mas$

The $AMT$ and $background$ form the two between-participant independent variables and the design can be described as: $3 (test) \times 2 (AMT) \times 2 (background)$ yielding 12 different conditions. A two-way repeated-measures ANOVA test was used to determine the significance of the results.

The same scenarios were used and all participants were asked the same questions in the same sequence. Collectively, these formed the control variables.

B. Materials

Each test utilised a corresponding attack scenario which was converted into an attack graph and corresponding fault tree. Attack scenario 1 (Figure 1) and 2 (Figure 2) were based on the fictional attack graphs produced by Barik and Mazumdar [47] and Ghosh and Ghosh [75] respectively. These scenarios were selected as they are published scenarios and have small visual structures in terms of the number of cyber-attack constructs being used.

Attack scenario 3 was developed by the authors and based on the Stuxnet attack - a well-known virus attack [107, 106]. The Stuxnet attack is very complex, and the resulting attack graph/fault tree contains more than sixty cyber-attack constructs. Consequently, a small section of the Stuxnet attack - representing the exploitation of the task scheduler vulnerability was used for the scenario.

Part of the scenario is presented in Figure 3. The comprehension context assumes that experts and non-experts are analysing and interpreting the visual syntax in particular - and not necessarily the formal syntactic definitions. In other words, the observer doesn’t necessarily need to have a technical understanding of the attack.

In a professional setting, although the analysis of an attack might begin by focussing on the full attack graph, quite often, the decision maker/analyst proceeds to focus on the poignant elements of the attack. Consequently, a small section of the overall graph was considered appropriate for the present study. A further exploration of the cognitive impact and effects of studying small/large attack models by different stakeholders over longer periods of time may be useful and is considered for future work.

The third scenario is quite different in style compared with scenarios 1 and 2. Scenario 3 is presented using a simpler explanatory narrative in comparison with scenarios 1 and 2. It was considered important to be able
Figure 1. Cyber-attack Scenario 1 represented as a fault tree (left) and an attack graph (right) (based on Barik and Mazumdar [47]). The three fundamental cyber-attack constructs are preconditions - represented as boxes in the attack graph and ellipses in the fault tree, exploits - represented as ellipses in the attack graph and boxes in the fault tree, and precondition logic - AND represented by the presence of an arc in the attack graph and 'semi-ellipse' in the fault tree and the OR represented by the absence of an arc in an attack graph and an 'elliptic triangle' in a fault tree.

Figure 2. Cyber-attack Scenario 2, fault tree (left), attack graph (right) (Based on Ghosh and Ghosh [75]).

Figure 3. A Section of Cyber-attack Scenario 3, fault tree (left), attack graph (right) based on a small part of the Stuxnet attack scenario described by Falliere et al., [106].
to represent and test a portion of a ‘real life’ attack to understand whether this impacted participant perception.

While the data content of the three scenarios differs (formal syntax versus narrativised representation), the aim of the study is to assess the cognitive impact of the visual syntax. Consequently, one might assume that the oth group are more likely to score higher - in test 3 in comparison with test 1 and 2, when presented with such a representation. The results (Table V) indicate no statistical differences in the ability of non-experts to perceive the attack descriptions either when presented in formal syntactic or textually narrativised terms.

The study was configured using the Qualtrics platform [111]. The study was divided into three sections which: gathered participant consent and other data; enabled the participant to gain fundamental background information relating to the AMT being studied; and then tested participant perception through a sequence of questions (Figure 4).

C. Participants

84 participants were recruited for the study. 21 participants did not complete the study leaving 63 participants. 43 males and 20 females aged between 21 and 58 (M = 29) were collected and grouped according to their background groups (cs n=31, oth n=32) and then further subdivided into AMT groups (aag n=31, ft n=32).

Participants were assigned to the aag/ft group randomly to avoid bias within these groups - particularly to avoid a situation where one group might understand cyber-attacks at a conceptual level better than the other.

D. Procedure

Participants were required to access the study by following a url. The experiment sequence was as follows:

a. The first screen provided general data regarding the study such as participant information, and required the participant to provide consent
b. The second screen gathered participant data (age, gender, experience)
c. The third screen enabled participants to study the ontology and structure of the AMT they had been assigned to.d. Following this, participants were required to complete the test

Each question in the test required the participant to study an AMT and answer a question. Both the aag and ft group were provided with exactly the same questions. Participants were not able to revisit questions.

V. RESULTS

Due to the large amount of data collected, the results are divided into subsections with a discussion following in section VI. The mean assessment scores for each test (mas1, mas2, mas3) were normalised and are outlined in Table IV. Figure 5 shows the mean assessment score (mas) vs group and mas vs AMT.

Table V highlights the mean differences between the groups. The mean difference is presented as a delta

\[ \delta \text{ value: } \delta_{\text{mas}(i)}_{\text{aag,ft}} = (\text{mas}(i)_{\text{aag}} - \text{mas}(i)_{\text{ft}}) \]

The mean differences for all three tests favours the aag group (\( \delta_{\text{mas} 1\text{aag,ft}} = 0.1463, \delta_{\text{mas} 2\text{aag,ft}} = 0.2475, \delta_{\text{mas} 3\text{aag,ft}} = 0.0406 \)). Furthermore the table outlines the mean differences for the overall tests - giving an overall \( \delta(\text{mas})_{\text{aag,ft}} = 0.1448 \).

These descriptive results suggest that the aag method is better at aiding attack perception when compared with the fault tree method. Whilst the mas and the mean differences appear to be significant, they require further investigation to determine statistical significance.

A. Main Effects

The main effects of test, AMT (aag/ft) and background (cs/oth) were analysed using a 3 (test) \( \times \) 2 (AMT) \( \times \) 2 (background) mixed design factorial ANOVA for all 63 participants as outlined in section IV-A. Mauchly’s Test of Sphericity indicated that the assumption of sphericity was not violated (Mauchly’s test, \( \chi^2(2) = 1.309, p = 0.520 \)).

1) Within-Participant Main Effects of test: The withinparticipants main effect for test signified that participant performance differed between tests (\( F(2, 118) = 3.232, p = 0.043 \)). Pair-wise comparisons were performed between the three tests and statistically significant results from this analysis are presented in Table VI and discussed further for the AMT (section V-B1) and background (section V-B2) groups.

2) Between-participants main effect of AMT: Hypothesis 1 (H1a) proposed that: the selection of AMT influences response to mas. Hypothesis 1 holds and the between-participants main effect for AMT revealed that the selection of AMT was significant, (\( F(1, 59) = 8.004, p=0.006 \)) showing that there is a distinction between the two AMTs with aag ranked higher than ft.

Table IV and Figure 5 outline differences in the performance between the aag/ft groups - favouring the aag group. (\( \delta_{\text{mas} 1\text{aag,ft}} = 0.1463, \delta_{\text{mas} 2\text{aag,ft}} = 0.2475, \delta_{\text{mas} 3\text{aag,ft}} = 0.0406 \)). The aag group demonstrated a better result overall for the test: \( \delta_{\text{mas}_{\text{aag,ft}}} = 0.1448 \). This is analysed further in section V-B1 to identify differences within the aag and ft groups.

3) Between-participants main effect of background: Hypothesis 2 (H2a) proposed that: the selection of background influences response to mas. Hypothesis 2 holds and the between-participants main effect for background was statistically significant (\( F(1, 59) = 12.843, p = 0.001 \)) showing that the selection of background influences response to mas with cs ranked higher than oth.
Table IV: Mean Assessment Scores mas

<table>
<thead>
<tr>
<th>Test</th>
<th>mas¹</th>
<th></th>
<th>mas²</th>
<th></th>
<th>mas³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>n</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>oth</td>
<td>aag</td>
<td>0.8353</td>
<td>0.2486</td>
<td>17</td>
<td>0.7582</td>
</tr>
<tr>
<td></td>
<td>ft</td>
<td>0.6267</td>
<td>0.3530</td>
<td>15</td>
<td>0.5020</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.7375</td>
<td>0.3152</td>
<td>32</td>
<td>0.6381</td>
</tr>
<tr>
<td>cs</td>
<td>aag</td>
<td>0.9357</td>
<td>0.1292</td>
<td>14</td>
<td>0.9386</td>
</tr>
<tr>
<td></td>
<td>ft</td>
<td>0.8294</td>
<td>0.2024</td>
<td>17</td>
<td>0.6718</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.8774</td>
<td>0.1788</td>
<td>31</td>
<td>0.7923</td>
</tr>
<tr>
<td></td>
<td>aag</td>
<td>0.8806</td>
<td>0.2068</td>
<td>31</td>
<td>0.8397</td>
</tr>
<tr>
<td></td>
<td>ft</td>
<td>0.7344</td>
<td>0.2966</td>
<td>32</td>
<td>0.5922</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>0.8063</td>
<td>0.2648</td>
<td>63</td>
<td>0.7140</td>
</tr>
</tbody>
</table>

δ_{aag:ft} = 0.1463, δ_{cs:aag:ft} = 0.2086, δ_{oth:aag:ft} = -0.0406

NB: δ values are calculated as absolute differences

Table V: Mean Differences by AMT and group

<table>
<thead>
<tr>
<th>Delta</th>
<th>mas¹</th>
<th></th>
<th>mas²</th>
<th></th>
<th>mas³</th>
<th></th>
<th>mas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>n</td>
<td>Mean</td>
<td>n</td>
<td>Mean</td>
<td>n</td>
<td>Mean</td>
</tr>
<tr>
<td>δ_{aag:ft}</td>
<td>0.1463</td>
<td>63</td>
<td>0.2475</td>
<td>63</td>
<td>0.0406</td>
<td>63</td>
<td>0.1448</td>
</tr>
<tr>
<td>δ_{cs:ft}</td>
<td>0.1063</td>
<td>31</td>
<td>0.2668</td>
<td>31</td>
<td>-0.0158</td>
<td>31</td>
<td>0.1191</td>
</tr>
<tr>
<td>δ_{oth:ft}</td>
<td>0.2086</td>
<td>32</td>
<td>0.2562</td>
<td>32</td>
<td>0.1392</td>
<td>32</td>
<td>0.2013</td>
</tr>
<tr>
<td>δ_{aag:oth}</td>
<td>0.1004</td>
<td>31</td>
<td>0.1804</td>
<td>31</td>
<td>0.2048</td>
<td>31</td>
<td>0.1619</td>
</tr>
<tr>
<td>δ_{ft:oth}</td>
<td>0.2027</td>
<td>32</td>
<td>0.1698</td>
<td>32</td>
<td>0.3598</td>
<td>32</td>
<td>0.2441</td>
</tr>
</tbody>
</table>

Figure 5. Overall means of test vs AMT (left) and test vs group (right)

Table V and Figure 5 highlight the mean assessment scores for the cs/oth groups. The Table shows that δ_{oth:(mas)_{aag:ft}} = 0.2013 and δ_{cs:(mas)_{aag:ft}} = 0.1191. The data outlines what appears to be a distinct difference favouring aag for both the cs and oth groups. This is analysed further in section V-B2 to determine if there is a statistically significant difference within the oth and cs groups.

B. Results for test, background and AMT

The main effects were analysed further to discover whether there are differences within the background and AMT groups.

1) Results for AMT: The main effect for test was further analysed using two: 3 (test) × 2 (background) ANOVAs for aag (n=32) and ft (n=31) to explore further effects. The between-participant effect signified that the selection of background influences response to mas (F(1, 61)=10.321, p=0.002).

a) Results for aag: The between-participants effect for AMT shows a statistically significant effect favouring cs (F(1, 29) = 6.396, p = 0.017) signifying that selection of background influences the mas score for the aag method. Post hoc tests using the Bonferroni correction showed that the cs group performed better than the oth group (mean = 0.162, SD=0.064, p=0.017). This is indicated by the overall mas results (Table V) which highlight...
a mas score favouring the cs group when using the attack graph method ($\delta_{cag_{cs,oth}} = 0.1619$).

Pair-wise comparisons were performed between the three tests for the aag group. There were no significant differences between the tests for the aag group.

b) Results for ft: Similarly, the between-participants effect for AMT shows a statistically significant effect favouring cs ($F(1,30) = 6.954, p = 0.013$) signifying that selection of background influences the mas score for the fault tree method. Post hoc tests using the Bonferroni correction showed that the cs group performed better than the oth group (mean $= 0.244$, $SD = 0.093$, $p = 0.013$). This is further indicated by the overall mas results (Table V) which highlights a mas score favouring the cs group when using the fault tree method ($\delta_{f\text{case,oth}} = 0.2441$).

Pair-wise comparisons were performed between the three tests between the ft group. Table VI shows that the ft group demonstrated a reduction in mas of 0.141 between tests 1 and 2 ($p = 0.032$) and an increase of 0.160 between 2 and 3 ($p = 0.009$).

2) Results for background: The main effect for test was further analysed using two: 3 (test) $\times$ 2 (AMT) ANOVAs for cs and oth to explore further effects. The between-participant effect signified that the selection of AMT influences the mas score ($F(1,61) = 5.536, p = 0.022$).

a) Results for cs: The between-participants effect for background shows a statistically significant effect for AMT ($F(1,29) = 5.276, p = 0.029$) signifying that the selection of AMT influences the mas score for the cs ft group. Post hoc tests using the Bonferroni correction showed the cs aag group performed better than the cs ft group (mean $= 0.119$, $SD = 0.052$, $p = 0.029$). This is further indicated by the overall mas results (Table V) which shows an overall mas favouring the cs aag ($\delta_{c\text{saag}_{ft}} = 0.1191$).

Pair-wise comparisons between the three tests (Table VI) show that the cs group demonstrated an increase in mas of 0.113 between 2 and 3 ($p = 0.029$).

b) Results for oth: The between-participants effect for background did not signify a statistically significant effect for AMT ($F(1,30) = 4.104, p = 0.052$) showing that the selection of AMT is not significant for participants in the oth group.

Pair-wise comparisons between the three tests for the oth group (Table VI) showed no significant differences between the mas scores for the oth group.

### Table VI: Statistically Significant Within Participants Effect Results for Tests

<table>
<thead>
<tr>
<th>Group</th>
<th>(i) test</th>
<th>(j) test</th>
<th>Mean difference (i-j)</th>
<th>Std error</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>all</td>
<td>1</td>
<td>2</td>
<td>0.089 (↓)</td>
<td>0.035</td>
<td>0.039</td>
</tr>
<tr>
<td>ft</td>
<td>1</td>
<td>2</td>
<td>0.141 (↓)</td>
<td>0.052</td>
<td>0.032</td>
</tr>
<tr>
<td>cs</td>
<td>2</td>
<td>3</td>
<td>0.113 (↑)</td>
<td>0.041</td>
<td>0.029</td>
</tr>
</tbody>
</table>

VI. DISCUSSION

This study investigated which of the two AMTs was more effective in aiding cyber-attack perception. The study also explored whether any demonstrable benefits of either of the two methods occurred under specific background grouping conditions.

Hypothesis 1 ($H1_1$) tested whether the selection of AMT influences response to mas. This hypothesis holds and the study finds that the aag method is better at aiding attack perception when compared with the fault tree method ($p < 0.01$). A straightforward comparison of the means for both groups reveals a mas favouring the aag method for every test.

Hypothesis 2 ($H2_1$) tested whether the selection of background influences response to mas. This hypothesis holds and the study finds that the aag method can be an effective tool for aiding cyber-attack perception amongst experts. As expected, participants that have a computer science background outperformed other participants when using both methods ($p < 0.05$) but demonstrated a preference for the aag method.

A comparison of the mas scores (Table IV) indicates a preference for the aag method compared with the fault tree method for the non-computer science group. However, this result was not statistically significant and has to be taken with caution.

Given the three tests tested progressively advanced cognitive levels - rather than testing the same cognitive level each time, one might expect a minimum desirable result to be a stable performance across the cognitive levels. However, differences in performance (Table VI) were demonstrated between some of the tests. This was most significant for the fault tree group between tests 1 and 2 (a decrease) and tests 2 and 3 (an increase). The performance increased between tests 2 and 3 for the computer science fault tree participants. There was no significant difference in the performance of the attack graph group. Notably, there was a consistent cognitive performance across the three tests for the aag method group with no significant decrease in performance.

Added to this, one might expect a better cognitive performance for the oth group for test 3 in comparison with tests 1 and 2 - given the difference in narrative style of the question. This was not the case, and as Table VI shows, there was no difference between the tests for the oth group.

Although the aag method rendered a better mas overall, there was a negligible difference in terms of mas3 which tests the ability to apply the two methods. In other words, participants in both the aag and ft were able to effectively apply the methods in given scenarios. This requires further research in a more focussed study.

VII. CONCLUSIONS

The growing number of cyber-attacks and the increased need for both experts and non-experts to better understand cyber-attack leads to the requirement for better techniques and methodologies that can be used to more quickly and effectively to appraise audiences on the methods used to perpetrate cyber-attacks and the weaknesses in systems that enable such attacks to prevail.

The main contribution of this research has been to show that the aag method is better than the fault tree method in aiding cyber-attack perception.
A. Limitations and Future Work

The present research has revealed limitations and areas for further research that should be explored to further reason with the results presented herein.

a) Benefits on perception of non-experts: The study revealed that the aag method could be suitable for aiding cyber-attack perception amongst non-expert audiences - such as CEOs. Although the results in this regard were not statistically significant, this requires a further larger study to identify whether the method can indeed aid cyber-attack perception amongst non-experts.

b) Measurement of Effectiveness: The present study assessed the ability of participants to comprehended the visual syntax. The study measured effectiveness as the ability to respond correctly to a question requiring the interpretation of the visual syntax of a given AMT. This definition can be expanded further to include timeliness and attack severity.

Although the time taken to complete the study was measured, this was not analysed in the measurement of ‘effectiveness’ because the present study considered it more important to allow participants to apply due consideration to each question than to place the participant under time based pressure.

An alternative analysis would be to take into account the correctness of the response as a function of time. The severity of an incorrect response could also be considered important. Further research requires the development of a methodology reflecting effectiveness based on three variables: correctness, time and severity.

c) Understanding key visual syntactic factors: Further research should attempt to understand which elements of the visual structure of an AMT - described in section III-C2, are more significant in aiding cyber-attack perception. In addition, it would be beneficial to understand the effect on perception of visual structural elements such as colour, tone, line width/density/structure. In particular, further research should explore improvements to aag which balance the tradeoff between providing more visual information - such as the inclusion of elements such as attacker capability and/or uncertainty, whilst maintaining effective cognitive perception.

The symbol count, (including a specific symbol for OR in an FT increases the total symbol count), the direction of information flow are likely to have impacted the results. In a follow up study, we will be comparing these in a subjective evaluation.

d) Acceptability amongst practitioners: It is critical that any method of representing cyber-attacks gains acceptance amongst practitioners. Such an audience would include teacher-practitioners (lecturers who might use the method to teach cyber-attack), non-expert corporate/decision makers and cyber-security analysts. Further research should test the aag method with these audiences to determine the efficacy of the method in aiding cyber-attack perception in a live environment.

e) Complexity: The attack scenarios used in the study were relatively small and comprised of up to 14 symbols (Figure 2). This is not representative of complex cyber-attacks such as the Stuxnet virus [107, 106], Jeep Cherokee Hack [112] and the Sony Hack [113, 114]. Further research should examine the effectiveness of the aag method when presenting larger scenarios.

ACKNOWLEDGMENTS

Kurt Debbatista was partially funded by a Royal Society Industrial Fellowship (IF130053)

REFERENCES


files/Root_Cause_Corrective%20Action.pdf

Author Biographies

Harjinder Singh Lallie holds an MSc in Information Networks, an MPhil (University of Birmingham) and a BSc in Industrial Information Technology. Harjinder is an experienced academic with more than 20 years of teaching and tutoring experience in the areas of cyber-security, computer networks and digital forensics.

Harjinder is the course leader for the MSc Cyber Security and Management degree. He teaches Digital Forensics at the University of Warwick and the University of Oxford (Centre for Doctoral Training in Cyber Security).

Harjinder has published in numerous peer-reviewed journals and conference and has sat on numerous research programme committees. Harjinder’s research interests include threat modeling, digital forensics and the presentation of digital evidence.

Kurt Debattista is an Associate Professor at the University of Warwick. His research interests include high-fidelity rendering, perceptual imaging, high dynamic range imaging, and high-performance computing. Debattista has a PhD in Computer Science from the University of Bristol, an MSc in Computer Science and an MSc in Psychology.

Jay Bal holds a PhD from the University of Warwick (WMG), an MSc in Integrated IC Systems Design (UMIST), and a BSc in Electronics and Management.

Jay has worked with GEC Telecommunications as a electronic circuit techniques designer, GEC Private Systems, Plessey Research and Rover Advanced Technology Centre.

Jay’s research interests include: virtual business ecosystems, virtual organisation breeding environments, creating value through the internet, electronic markets and improving the success rate of IT projects.