Why People Still Fall for Phishing Emails: An Empirical Investigation into How Users Make Email Response Decisions

Asangi Jayatilaka*†, Nalin Asanka Gamagedara Arachchilage‡, Muhammad Ali Babar*
*Centre for Research on Engineering Software Technologies (CREST), The University of Adelaide, Australia
{asangi.jayatilaka, ali.babar}@adelaide.edu.au
†School of Computing Technologies, RMIT University, Australia
asangi.jayatilaka@rmit.edu.au
‡School of Computer Science, The University of Auckland, New Zealand
nalin.arachchilage@auckland.ac.nz

***Abstract*—Despite technical and non-technical countermeasures, humans continue to be tricked by phishing emails. How users make email response decisions is a missing piece in the puzzle to identifying why people still fall for phishing emails. We conducted an empirical study using a think-aloud method to investigate how people make 'response decisions' while reading emails. The grounded theory analysis of the in-depth qualitative data has enabled us to identify different elements of email users' decision-making that influence their email response decisions. Furthermore, we developed a theoretical model that explains how people could be driven to respond to emails based on the identified elements of users' email decision-making processes and the relationships uncovered from the data. The findings provide deeper insights into phishing email susceptibility due to people's email response decision-making behavior. We also discuss the implications of our findings for designers and researchers working in anti-phishing training, education, and awareness interventions.**

I. Introduction

Phishing is one of the most prominent and influential cyber attacks, as it is a precursor for many other attacks, including identity theft, which is among the worst [16], [28]. Phishing attacks have sharply risen recently, partly driven by COVID-19 and supply chain uncertainty [1]. Phishing activity trends report of the Anti-Phishing Working Group reports more than 4.7 million attacks for 2022. This report also highlights that since the beginning of 2019, the number of phishing attacks has grown by more than 150% per year. Phishing attacks are most commonly launched through emails as they are difficult to detect [60], [70]. Often phishing email attacks target banks, defense organizations, and private companies, as they curate a wide variety of data, including personal and financial data [9]. Generally, clicking on links, downloading attachments, or replying to phishing emails can be considered unsafe response decisions [39], [54]. CISCO's 2021 Cybersecurity threat trends report states that at least one person has clicked a phishing link in around 86% of organizations [15]. A successful phishing attack can trick users into unintentionally disclosing their valuable information, compromising their devices or accounts [67]. Additionally, phishing attacks are also used for installing malware (i.e., malicious software), which can disturb the normal operations of a computer system, contributing to significant financial losses and reputation damages.

Technology alone is insufficient to combat phishing email attacks; therefore, transforming users from the weakest line of defense to the most robust line of defense is essential [21], [29], [57]. An active community of practitioners and researchers focuses on phishing education, training, and awareness (e.g., phishing alerts) to support users in correctly identifying phishing [18]. However, these efforts are with limited success [4], [21]. A major challenge in designing effective anti-phishing interventions is the lack of attention to reasons why people still fall for phishing [21], [37], [50], [75].

Understanding why people still fall for phishing emails will provide underpinning science to design effective future anti-phishing tools and educational interventions. Although prior literature largely focuses on analyzing the personality and demographics of people who fall for phishing emails [43], [66], how people make email responses and the thought process a user goes through when deciding how to respond to their emails is often overlooked [21], [34], [37], [50], [75]. Only a limited number of studies have attempted to conduct qualitative user studies to explain people's email decision-making processes [34], [57], [75]–[77]. Although qualitative studies, compared to quantitative studies, allow us to obtain more detailed and holistic insights into users' email response decision-making behaviors and the reasons for those behaviors, even such prior work does not interpret the different elements of people's email response decision-making processes and their relationships influencing their email response behavior.

To address this research gap, we conducted an empirical investigation through a "think-aloud" role-play experiment and follow-up interviews to better understand people's decision-making behavior when responding to emails. We developed a theoretical model that explains how people are driven to respond to emails by clicking on links, replying, and downloading attachments based on the identified elements of the

Symposium on Usable Security and Privacy (USEC) 2024
26 February 2024, San Diego, CA, USA
ISBN 979-8-9894372-5-2
https://dx.doi.org/10.14722/usec.2024.23072
www.ndss-symposium.org, https://www.usablesecurity.net/USEC/

people's email response decision-making process and the relationships uncovered from collected data. The model developed based on empirical evidence interprets how different elements of people's email response decision-making processes could positively and negatively influence people's intention to respond to emails, which was lacking in previous literature. For example, the model provides deeper insights into how certain habits, validation techniques, previous experiences, etc., can positively influence people's intention to respond to emails, as a result increasing the risk of them falling for potential phishing attacks. In summary, our contributions are as follows:

- Based on empirical evidence and grounded theory analysis, we provide knowledge into elements of email users' decision-making process (e.g., diverse types of emotions, personal habits, previous experiences, and email characteristics) that influence their email response decisions.
- We develop a theoretical model that explains how people are driven to respond to emails by clicking on links, replying, and downloading attachments based on the identified elements of the user's email response decision-making process and their relationships uncovered from data. The developed theoretical model provides deep insights (i.e., scientific underpinnings) into an individual's email decision-making process or response behaviors to emails in general. As a result, the model enables us to identify general email decision-making flaws that attackers could potentially exploit to launch successful phishing attacks. Furthermore, understanding people's general email decision-making flaws the attacker could leverage may help better design technical countermeasures such as anti-phishing tools to thwart email-based phishing attacks.

II. Background and motivation

It is imperative to obtain a deeper understanding of the scientific underpinnings of how users make email response decisions to design better technical and non-technical countermeasures that thwart email-based phishing attacks. Research on phishing email susceptibility has considered the demographic or personality of victims [23], [36], [39], or phishing email characteristics [12], [20], [43], [81]. Conversely, several studies have investigated psychological and behavioral responses in this regard [26], [54], [81]; they mostly employ existing theories borrowed from other fields to derive phishing models [25], [38], [48], [49], [65], [72]–[74]. Such work often employs hypothesis-testing to see if the theory, specified as the hypothesis, is supported by gathered quantitative data.

While models derived from existing theories and validated with quantitative data provide much-needed insights into phishing, they often focus only on specifically selected dimensions of email response behaviors depending on the selected theories. For example, researchers in [48] focused on the phishing cues available in the emails. They hypothesized that phishing cues are linearly combinable and hence a type of "Judgment Analysis", is appropriate for evaluating phishing judgments. Through role-play, experiments conducted with participants who judged whether emails were phishing proved that their hypothesis was correct. In another study [25], researchers focused on the impact of the content and framing of phishing emails on user vulnerability. They came up with several hypotheses and evaluated those using surveys with University students. Their results suggest that the desire to protect things of value and the opportunity to obtain valued objects could make people vulnerable to phishing attacks. Researchers in [65] derived several hypotheses based on the protection motivation theory and the theory of planned behavior, focusing on the individual, organizational, and technological factors that affect phishing email responses. They later tested the validity of the derived hypothesis using questionnaires. Inspired by research on information process and interpersonal deception, researchers in [72] developed an integrated information processing model of phishing vulnerability. Upon validating the model with undergraduates, the researchers found attention to the email source, grammar and spelling, urgency cues, and subject line were significantly negatively related to an individual's likelihood of responding to a phishing email.

Only a limited number of studies have conducted qualitative user studies to propose theoretical models to explain people's email decision-making processes [34], [75]–[77]. For example, Wash [75] conducted interviews with experts to identify the process that they follow to identify phishing messages they received in the past. They found that experts follow a three-stage process for identifying phishing emails — first sense-making, then suspicion, then acting. On the other hand, Jayatilaka et al. [34] revealed eleven high-level themes influencing people's email response behaviors [34]. However, the study [34] failed to interpret the elements that constitute these themes and the nature of their relationships leading to the email responses.

In addition to the above-mentioned gaps, several other common drawbacks that exist in most aforementioned studies are described below. Firstly, it is common for images of emails to be used in phishing-related experiments [17], [59], [83]; however, this could detach participants from their naturalistic setting, affecting their decision-making. Moreover, such images prevent understanding whether or not people use the link URL information shown in the status bar without alerting participants [54], [73]. Secondly, surveys are often utilized to collect data from the users about their habits, traits, and explanations for decisions they have made [2], [14], [32], [51], [72], [74], [74], [77]. However, such self-reported surveys may not provide adequate information or be inaccurate as users could justify the corresponding email response. Jaeger et al. [33] have used eye-tracking software in their experiment to reduce the limitations of using a survey. Thirdly, most research request participants to make pre-defined email legitimacy judgments or response decisions (e.g., phishing or legitimate) [11], [36], [40], [59], [61], [73], [79], [83]; however, this may not align with how people behave naturally. For example, our study findings reveal that, at times, people could develop doubts about email legitimacy and hence make a final decision only after validating the email. Fourthly, several research studies have relied on participants' descriptions of past incidents/situations to understand how they detect phishing emails [75], [77]. Such methods could suffer from the imperfect memory of participants of past incidents.

The current study aims to provide deeper insights into users' email response decision-making processes by devel-

oping a theoretical model explaining how people are driven to respond to emails by clicking on links, replying, and downloading attachments based on elements of the user's email response decision-making process and their positive and negative relationships uncovered from grounded theory analysis of qualitative data. To gather qualitative data about participants' email response decision-making behaviors we employed a role-play-based think-aloud method and follow-up interviews using a simulated email client. The reasons for this data collection method and set-up are further explained in Section III. Armed with this deeper understanding of email response decision-making elements and their relationships, we can begin to design more effective anti-phishing tools and techniques to face phishing attacks successfully.

III. Research Methodology

We decided to use a role-play-based think-aloud method to achieve the aim of this research instead of surveys [2], [14], [32], [51], [72], [74], [77] or retrospective interviewing [75], [77], as those: i) rely on memory of past incidents, ii) don't provide detailed insights into the decision-making processes, or iii) rely on people's justifications for email responses already determined. In our study, participants were not requested/forced to make certain decisions (e.g., classifying emails as legitimate/phishing), allowing them to explain their intention to respond to emails naturally. Having a simulated email client (see Figure 1) enables the participants to naturally engage with the activity while contextualizing their opinions and explaining how they would respond to emails by recalling their memories in a naturalistic environment. We conducted follow-up interviews to expand on the think-aloud results. As the emails were embedded in the simulated email client, we were able to automatically adjust the dates and times specified in the email content to suit the time and day that a participant carried out the experiment. The research was approved by the ethics committee of the University of Adelaide.

In this study, in terms of email responses, we focus on clicking on links, replying, and downloading attachments because, as mentioned before, they are considered unsafe email responses in the context of phishing. Henceforth, for brevity, we refer to clicking on links, replying, and downloading attachments as email responses throughout this paper.

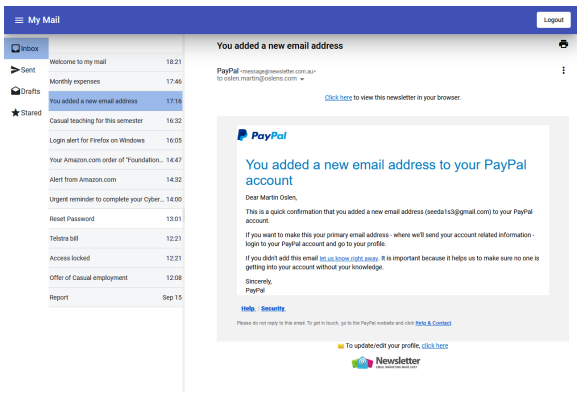

Fig. 1. Simulated web email client

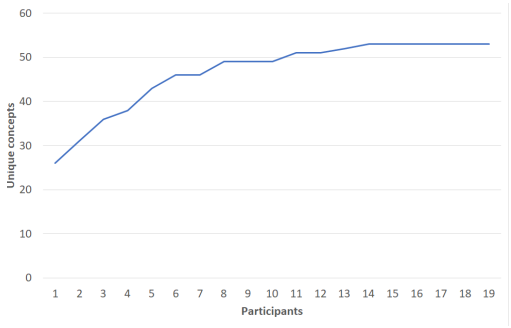

Fig. 2. The number of unique concepts for each participant

A. Emails selection

We used real phishing emails (12) and corresponding real legitimate emails (12) for this study[1]. As we aim to obtain a holistic picture of how people are driven to respond to emails, the selected emails reflect different domains (e.g., education, financial, social media), and attacker strategies. Examples of emails include University password expiry, Facebook alerts, Amazon order confirmations, and mobile bills. Several phishing strategies, such as providing a personal salutation, creating fear and urgency, mimicking the appearance of a legitimate email, mimicking the sender's email address, including the signature of the sender, and URL obfuscation, highlighted in the literature that is used to to create fraudulent emails were included in the selected list of emails [47], [68]. The phishing emails were sourced from various venues [8], [62], [64]. The emails were adapted to the scenario (e.g., the name of the recipient was changed to the name of the fictitious character in the role-play). The corresponding legitimate emails were sourced from researchers' email correspondences (e.g., if the phishing email is about a shared Google spreadsheet, the corresponding legitimate email was also chosen to be about a shared Google spreadsheet). Based on the email assignment process which is described at the end of this sub-section, a participant did not receive a legitimate email and its corresponding phishing email together in his or her inbox.

Real phishing links from PhishTank [56] were adapted as phishing URLs for this study. Recent research points to six URL obfuscation techniques [19], [22]. We included two phishing emails for all except one URL obfuscation technique (i.e., obfuscating with HTTPS schema) pointed out [19]. For obfuscating with HTTPS Schema to occur, the browser must navigate to the destination (i.e., a landing page) to show the green padlock icon on the address bar. However, as previous studies have indicated, most people clicking on phishing links will go on to disclose information to those phishing websites [66]. Therefore, as our work focuses on understanding how people make email response decisions while reading their emails, the browser-based email client was prevented from navigating to any URL destination when a participant clicked on an email link. Instead, regardless of its legitimacy, a "link clicked" message was shown to the users when they clicked on any email link. Therefore, the obfuscation technique with HTTPS Schema is not meaningful in our study [19]. Most

[1]Screenshots of the phishing and legitimate emails used for the study are included in this link: https://osf.io/gu7qs/?view_only=1aa1e0ab5e47440aa5cb4e7a6a49d9c8

research on phishing emails focuses on emails with links neglecting emails without links; however, as explained before, phishing emails can come as emails without links. Therefore, we included several emails without any links. This included two emails requesting download attachments and two emails requesting to reply.

When assigning the emails, we selected 12 emails from the pool of emails. Then, considering a phishing email and the corresponding legitimate email as a pair, we randomly assigned the legitimate email to 50% of the participants and the phishing email to the remaining 50%. Following this approach we ensured that a given participant will not receive a legitimate email and its corresponding phishing email. Also, given an email pair, 50% of the participants will receive legitimate, and the remaining 50% will receive the corresponding phishing email to reduce any biasness in the analysis.

B. Data collection

We recruited 19 students from the University of Adelaide by distributing flyers. Eight participants were male and 11 were female. Twelve were undergraduate students and 7 were post-graduate students. They were distributed across the Faculty of Arts (5), Faculty of Health and Medical Sciences (5), Faculty of the Professions (4), and Faculty of STEM (5). Fourteen participants were 25 years of age and younger, and 5 were over 25. Six participants mentioned that they have had some form of anti-phishing training. Out of these six participants, 4 were undergraduates and 2 were post-graduates. Each data collection session was conducted over Zoom and lasted for around 90 minutes. The participants were requested to share their screens with the researcher so that the researcher was able to better understand how the participants interacted with the emails. These Zoom sessions were recorded and later used to generate the transcripts for the analysis.

Participants were asked to assume they were a fictitious person named "Martin Oslen". The background information on Martin was provided to the participants through a document. The fictitious person ("Martin") attends the participants' University, and the scenario included information on Martin's: i) social and professional interactions with their contacts (e.g., supervisor details, the mobile service provider, wife's details); and ii) online services and platforms (e.g., details about his social media, platforms he uses for online shopping such as Amazon and his banking details). We asked participants to role-play and debrief Martin's scenario before opening his email inbox. We allowed them to retain and refer to the document whenever they needed to throughout the study. When the participants were comfortable with the scenario they were given the URL of the email client and were asked to share their screens. We then provided basic descriptions of the email client to the participants. Furthermore, before the think-aloud session, participants were given a practice email to get familiar with the setup.

During the think-aloud session, we requested the participants explain how they felt to receive each of the given emails and how they would react to each one of them. We asked participants to talk freely about anything that came to their minds when going through the emails. We explained to them that the goal of the study is to obtain deeper insights into how they make email responses usually and not to evaluate the appropriateness of those decisions or responses. Therefore, they were requested to act as they would normally do for the given emails. In this process, we made sure not to request participants to perform certain actions (e.g., instructing participants to hover their mouse over a link to see the destination URL in the status bar [30]) or make certain decisions (e.g., classifying whether an email is legitimate or phishing) as that could influence their usual behavior. The follow-up questions were asked based on what participants explained in the think-aloud session to further clarify their thought processes.

Two pilots were conducted which allowed the researchers to identify and introduce three changes to how the think-aloud sessions and the follow-up interviews were conducted. During the first pilot, we identified that the initial question we asked during the think-aloud session (i.e., "how legitimate do they think a given email is") restricted the participants from showing their natural behaviors. That question forced them to first make a decision on the email's legitimacy without much consideration and later think of justifications to support their decisions. We decided to replace this question during the second pilot session with "how do you feel about receiving this email". This allowed the participants to behave more naturally and openly discuss what came to their minds and later explain the desired actions. Secondly, the pilot interviews also demonstrated the importance of providing a practice email to participants so they can be familiarized with the setup. Thirdly, we asked follow-up questions after each email in the second pilot as participants faced difficulties answering the questions when follow-up questions were asked at the end of all the emails in the first pilot. We observed that the approach used in the second pilot was more convenient for the participants to answer the questions while their memory was still fresh.

C. Data analysis

We used Grounded Theory (GT) [24], [69] for data analysis, using NVivo^{TM} software. The data analysis included different types of coding: ***open coding***, ***axial coding*** and ***selective coding*** [69]. During ***open coding***, we read through the interview transcripts line by line and encapsulated them into codes with short phrases. Later we performed a second round of open coding to refine the results of the first round. We coded the interviews based on contextualized statements instead of single terms. After coding, we discussed the relations between the newly found codes and agreed upon a set of higher-level axial codes (i.e., ***axial coding*** and these will be called concepts hereafter). The concepts served as a baseline for the next round of coding, which is ***selective coding***. During the selective coding process, all researchers agreed upon a set of codes (called core categories hereafter) that represent the different elements that affect the employees' email response behaviors. The first author led the data analysis, and the second author reviewed the emerging codes, concepts, and categories along with the interview transcripts during each step of the data analysis process; any disagreements were resolved through discussion before moving further into the analysis. Frequent iterative group meetings with all the authors were held throughout the process to ensure data interpretation consistency and rigor. All three authors have extensive experience with various qualitative methods including grounded theory and thematic analysis. We achieved saturation through this process (i.e.,

where no fresh information emerges from subsequent think-aloud sessions) because all the main categories and concepts had been uncovered across the 19 participants (see Figure 2).

The last step of our GT approach was to form a theoretical model by considering the relationships between the discovered axial and selective codes. Similar to a previous study [46], the model was based on the knowledge (i.e., categories and concepts) that emerged from grounded theory analysis. Model relationships were identified through rigorous and iterative analysis of transcripts based on "interpretive data" of people's email response behaviors. All researchers discussed the draft theoretical model to reach an agreement. We ensured the completeness and accuracy of the model in terms of the collected data through negative case analysis [10], [46]. Here we went through interviews iteratively to check whether the participants' statements could be assigned to the draft theoretical model. If not, we identified how they diverged from our draft and adopted it accordingly. We iteratively refined our theoretical model until all statements were captured.

It is important to note that, as the goal of the study was to explore individuals' decision-making processes that influence their email response decisions, all concepts and relationships that emerged from the qualitative data (involving 19 participants where each participant went through 12 emails during the study) are included in the model regardless of their frequencies. Therefore, not all categories and concepts are reflected in all our participants' email response decision-making processes. As quantitative results in qualitative research cannot be used to generalize findings; hence we will discuss all statements in Section IV without providing numbers.

IV. Results

An overview of the theoretical model that emerged from the data is presented in Figure 3. In summary, an individual's email response behavior is determined by how they perceived the email legitimacy, which in turn is influenced by the concepts of perceived sender legitimacy, perceived familiarity, perceived professionalism in the email title and body, perceived likelihood of receiving the email, perceived adequacy of length and granularity of information, trust for email links, perceived sense of security from auxiliary security content and previous phishing experiences. Furthermore, an individual's intention to respond to an email is also determined by their emotional attachments and personal habits, as well as their intention to validate the email. In-depth knowledge about diverse concepts that constitute the high-level categories and the nature of their relationships leading to email responses is provided below.

A. H1: Effects of "perceived sender legitimacy" on "perceived email legitimacy"

Our findings reveal that how people perceive the sender's legitimacy is pivotal in the trust they place in a received email. People tend to focus on specific components of the sender's email address, sometimes, without even having a clear understanding of which components need more attention. Given the sender's address, most participants fail to place more emphasis on the domain than the sub-domain, email display name, and sender's user name. One participant explained that he is convinced that a given email is coming from a reliable source by only considering the sender's email display name overlooking the issues he observed in the sender's domain "*It is a bit dodgy. But I'm convinced it's from PayPal [pointing to the sender's email name].*" (P15–P) [see H1.1]. Some participants made decisions about the sender's legitimacy by looking at the sender's user name. Having no-reply as the user name, having a familiar or professional user name allows a user to have more faith in the sender [see H1.2]. Also, participants focused on the domain of the sender address in deciding the sender's legitimacy [see H1.3]. We often notice when the sender address domain is deemed to be known or looked professional, participants tend to believe the sender is legitimate. In some instances, we observed that participants get confused with the sub-domain and domain. They, at times, cannot differentiate the domain and sub-domain. A participant explained that she is not trusting the sender as she was not confident about the sub-domain of the sender's email address even when she was confident about the domain "*I don't think that they have edm [pointing to the sub-domain] in the bank emails. It doesn't make sense*" (P02–L) [see H1.4].

A few participants looked into the reply-to-address to draw conclusions about the sender's legitimacy. For example, one participant got suspicious after seeing the reply-to address and identifying that the reply could be going to a Gmail account instead of the sender's official address "*When I reply, it will go to this Gmail, I would delete my email*" (P06–P). On the other hand, there was another participant who suspected the sender's address; however, after looking at the reply-to address (the user name of the reply-to address), she was much satisfied with the sender's legitimacy "*This email is there [mary@sau23.org in as the reply-to email]. Mary at dot org*" (P01–P) [see H1.5]. We also observed situations where participants had issues trusting the received emails after noticing the spelling issues in the sender's email address [see H1.6].

People also look at the sender's full email address to make conclusions about the sender's legitimacy. If the sender's address is known or familiar to the participants, then they seem to instantly develop trust about the sender "*That's like a staff email address from my experience. I feel good*" (P16–P) [see H1.7]. Some participants made assumptions about the sender based on the email addresses specified in the email body [see H1.8]. For example, a participant was satisfied with the sender's legitimacy after seeing the address specified in the body of a shared google sheet and overlooked the issues that he observed in the sender's email address specified in the email header "*It looks like just a shared thing from the wife [looking at the email adders in the email body]*" (P09–P). Therefore, based on the data, the following relationships exist:

H1.1 - H1.5: Perceived legitimacy of the [sender's email display name, sender address user name, sender address domain, sender address sub-domain, reply-to address] positively affects perceived email legitimacy
H1.6: Spelling issues in the sender email negatively affect perceived email legitimacy
H1.7 - H1.8: [Knownness of the sender address in the email header, knownness of the sender addresses specified in the email body] positively affects perceived email legitimacy

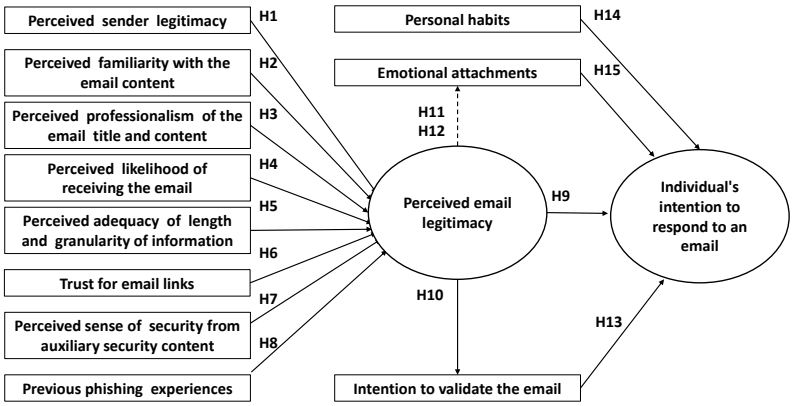

Fig. 3. Overview of the theoretical model developed based on gathered qualitative data. The latent variables are indicated with ovals, and core categories derived based on GT analysis are shown in rectangles. Concepts are not shown in the figure for brevity. In a 'partially affect' relationship only some concepts in a category are affected by the other category or latent variable.

B. H2: Effects of "perceived familiarity of the email content" on "perceived email legitimacy"

We observed people often intentionally and unintentionally compare the received emails with what they have seen before. More specifically, our results explain which aspect of emails people particularly consider when making judgments about the familiarity of the email content. Based on the data people consider the familiarity of the writing style [see H2.1], the familiarity of the interface elements [see H2.2], and the familiarity with email layouts [see H2.3] when making judgments about email legitimacy. For example, one participant explained her reasons for trusting a given phishing email "*It's similar to the emails that I know to be real emails from the University Human Resources. So, I already trusted the layout in my head*" (P05–P). There were situations where participants made decisions about email legitimacy only based on their perceived familiarity with writing styles, interface icons, or email layouts. For example, "*Okay, it's a Google Sheets looks very fine. I don't think I wouldn't check the sender's address in this one*" (P18–L). Hence based on our data, the following relationships exist:

> **H2.1 - H2.3:** Familiarity with the [writing style, interface elements, email layout] positively affects perceived email legitimacy

C. H3: Effects of "perceived professionalism of the email title and content" on "perceived email legitimacy"

Our data reveals insights into specific aspects (e.g., email layout [see H3.1], writing style [see H3.2], and interface elements [see H3.3]) people tend to focus in terms of professionalism in emails. With respect to the writing styles, participants looked into the professionalism in the writing styles of the email title and email body. For example, a participant explained that she is not trusting a given legitimate email because she is not convinced that the title of the email is professional enough "*Because it should have a kind of professional way like Notification of Declining Transaction something like that*" (P01–P). In terms of layout, participants considered how the email is organized, including the layout of the email links. One participant explained seeing other links such as 'contact us in an email makes them trust those emails "*They have these contact links, which would certainly be in a reliable source. So, I think including these links is really frequent in, you know, in a trusted company*" (P01–P). In terms of interface elements, participants explained that they look for professionalism in color schemes, buttons, font styles, and logos used in the email. There were instances where participants got suspicious about email legitimacy, even for a legitimate email, when they thought the email didn't have professional-looking interface elements, for example, buttons "*I don't look at anything else. After seeing that [button for the link], I am pretty much sure that it's a phishing email*" (P14–L). Therefore, based on our data, the following relationships exist:

> **H3.1 - H3.3:** Perceived professionalism in the [email layout, writing style, interface elements] positively affects perceived email legitimacy

D. H4: Effects of "perceived likelihood of receiving the email" on "perceived email legitimacy"

People often trust an email in situations where they are expecting such an email [see H4.1]. For example, a participant mentioned that they could be quite confident about an Amazon order confirmation email if he knew that he had placed such an order "*I think I would be very confident getting this email if I know that I have placed an order. I think this is kind of an expected email*" (P02–P). We also observed that participants even tend to ignore phishing cues they observe in emails if they perceive the likelihood of receiving those emails is high.

> **H4.1:** Perceived expectancy of the email positively affects perceived email legitimacy

E. H5: Effects of "perceived adequacy of length and granularity of information" on "perceived email legitimacy"

Our findings reveal that, at times, people tend to look into the length and granularity of the email when making email

legitimacy judgments. There seems to be a misconception that phishing attackers would not create lengthy emails or will not have the capacity to access their granular personal information (e.g., name, account numbers, etc). In terms of the length of an email [see H5.1], P1 explained that she does not think a phishing attacker would create long emails to make people fall for those emails, so she has much trust in lengthy emails "*This content being a really long email, I think they are actually trying to convey me this message, not a spam-like ... I think a spammer doesn't want to type this much to make a person fall into trap*" (P01–P). On the other hand, when the email has detailed information about the situation and/or information specific to the participants, they tend to believe it is a legitimate email [see H5.2]. For example, a participant explained, "*I think this is a legitimate email because it includes an item that I ordered and the approximate arriving date and the address. So they have the price. They have a lot of details that they should be actually knowing and that a spam email would not know.*" (P01–L) Therefore, based on our data, the following relationships exist:

H5.1 - H5.2: Perceived adequacy of the [email length, granularity of information] positively affects perceived email legitimacy

F. H6: Effects of "trust for email links" on "perceived email legitimacy"

Our results reveal that the absence of email links tends to create a false sense of security in email recipients [see H6.1] "*They don't want me to click on something right. I will just use regular reply to reply to this email*" (P18–P). Most of the time, our participants did not realize that emails without links could also be unsafe as there is a possibility of exposing their personal information by replying to those emails or spreading malware by downloading attachments in phishing emails. On the other hand, when there are links in emails, the trust for those links plays a pivotal role in people's intention to respond. Some participants explained that they would trust links that appear to be non-mandatory [see H6.2] "*It is not asking to do anything particularly. This creates a huge trust in me. It is not asking me to click a link*" (P19–P).

We also observed that participants come to conclusions about the trustworthiness of the email links based on the link text and appearance [see H6.3]. Some participants explained that they prefer to see the full URL specified in the email body rather than having an alternative text or an image, not realizing that the actual destination can be different from the URL text specified in the email body "*They are not showing the URL just have the word here [link text]. Why do they have to hide the URL. It's probably because it's a fake website*" (P04–P). Some looked at the button's appearance and made conclusions about the link's legitimacy. Some even believed that if a link text starts with HTTPS, it simply means a secure connection and the link is trustworthy "*They [links] usually start with HTTPS. If there's no HTTPS, I might be a little concerned*" (P18–L).

We only observed a few participants who determined the link destination correctly by hovering over or copying the link address. However, even they made incorrect judgments about link destinations as a result of a lack of understanding of URL structures. For example, some were satisfied seeing the organization's name anywhere in the destination URL. These findings are in line with previous work [3] on phishing URLs that explained that people are strongly biased towards answering that a URL would lead to the website of the organization whose name appeared in the URL, regardless of its position in the URL structure [see H6.4] "*I am just looking at the bottom left in the corner. As it has 'confirmation', 'account', and 'security', yeah, it does sound familiar*" (P18–P). Some even make decisions on the link legitimacy based on the similarity of the URL mentioned in the link text and the destination URL [see H6.5]. Some had doubts about the email links when they observed unfamiliar text or numbers in the destination URL [see H6.6] "*It [destination URL] looks really weird to me. Those numbers, I don't know what exactly that means. It makes me feel very uncomfortable clicking on it. Probably, it could contain a lot of information in those numbers. But I don't know what it is*" (P18–P). Therefore, based on our data, the following relationships exist:

H6.1 - H6.5: [Absence of email links, Perceived non-mandatory nature of the links, Perceived trust about links based on the link text and appearance, Appearance of the organization name in the destination URL, Similarity of the URL specified in the link text and the destination URL] positively affects perceived email legitimacy
H6.6: Appearance of unfamiliar text or numbers in the destination URL negatively affects perceived email legitimacy

G. H7: Effects of "perceived sense of security from auxiliary security content" on "perceived email legitimacy"

Our data revealed that people feel a sense of security when an email indicates that it has been scanned by an external scanning tool [see H7.1]. One participant explained that he feels he could trust emails that say external tools have scanned the content even when he does not understand the role of a scanning tool "*But yeah, this message here [message saying that the email is scanned for malware by force point] might probably make me loose guard a bit*" (P18–L). On the other hand, we observed that participants tend to trust emails after reading the information provided in the email footer [see H7.2]. For example, if the email footer provides reasons why they got the email or organization information such as ABN, or copyright information, then people tend to believe they have received a legitimate email "*Just having that claim at the bottom [footer] explaining why I got this email. That makes me confident enough to open up that straight away*" (P17–P). Some believe that information contained in the email footer about the company can be easily verified through the internet, not realizing that attackers could also obtain that information without much difficulty. Some even analyzed the phishing education, training, and awareness messages provided in the email [see H7.3]. They tend to believe that if an email provides users with anti-phishing education, there is less possibility for it to be a phishing attack "*Fake emails usually don't give you methods to identify fake emails*" (P18–L).

H7.1 - H7.3: Perceived sense of security based on [information on external scanning tools specified in the email, email footer information, in-email security education, and awareness] positively affects perceived email legitimacy

H. H8: Effects of "previous phishing experiences" on "perceived email legitimacy"

Our data reveal that past phishing email encounters could make people suspicious about future emails [see H8.1] "*No, I will never click email link because I have learned a lesson*" (P1–L). People can self-learn correct as well as incorrect strategies for detecting phishing emails based on phishing emails they have seen before [see H8.2]. For example, one participant explained his suspicion over a user name of the sender's email address as he has seen the same user name in phishing emails before "*I mean so far with my experience ... I have seen this webmaster, this kind of things. It is generally used by falsified email thing*" (P14–L). Several participants indicated that they had received some form of formal education related to phishing, and as a result, they are aware of some strategies to identify phishing emails [see H8.3] "*I also check the punctuation because I've learned from my computer class that phishing emails have misspellings and issues with punctuation*" (P01–P).

H8.1: Suspicion for email based on previous phishing email encounters negatively affects perceived email legitimacy
H8.2 - H8.3 Trust perceived by applying [self-learned correct and incorrect strategies from past phishing encounters, strategies learned from formal education and training] to detect phishing emails positively affects perceived email legitimacy

I. H9: Effects of "perceived email legitimacy" on "intention to respond an email"

Our data reveal that often participants decide to click on email links, reply to emails, or download attachments when they perceive an email as legitimate [see H9]. For example, a participant, after determining a given phishing email is legitimate, said he will reply to the email immediately as he trusts that email "*Yeah, this one's legitimate ... So I will reply like like Hi I'm Martin. so I'll be fine with this.*" (P13–P). Therefore, based on our data, the following relationship exist

H9: Perceived email legitimacy positively affects the intention to respond

J. H10: Effects of "perceived email legitimacy" on "intention to validate the email"

Our data provides evidence that when people have doubts about the email's legitimacy, they may want to validate the email further before deciding on the final response [see H10]. In this instance, their email response decision depends on how much trust they can develop about the email legitimacy based on the validation technique instead of the originally perceived email legitimacy "*I get a feeling that this might not be a true email. So, first, the action I will be taking is to go to the website and make a call to them*" (P01–P). Details on the types of validation techniques participants described and their effects on the intentions to respond to an email are described in Section IV-L. Therefore, based on our data, the following relationships exist:

H10: Perceived email legitimacy negatively affects the intention to validate the email

K. H11, H12: Effects of "perceived email legitimacy" on "emotional attachments"

We observed situations where participants get furious or frustrated after perceiving an email as phishing. These specific emotions will then drive their response behaviors later. For example, sometimes, participants expressed their desire to avoid the current email [see H11] as well as future emails [see H12] based on anger/frustration resulting from the identification of a phishing email. We provide detailed explanations on how these emotions drive their email response behaviors in Section IV-N. Therefore, based on our data, the following relationships exist

H11 - H12: Perceived email legitimacy negatively affects [motivation to avoid the current email based on anger/frustration resulting from the identification of a phishing email, motivation to avoid future phishing emails based on anger/frustration resulting from the identification of a phishing email]

L. H13: Effects of "intention to validate the email" on "intention to respond an email"

As explained in Section IV-J, our findings reveal that email validation is triggered by doubts about email legitimacy. The trust perceived after validation positively influences the intention to respond. Unfortunately, our data revealed that while some of the email validation techniques people use are safe, others are unsafe.

Some participants search the internet to look for information and logos specified in the email [see H13.1]. They seem to be satisfied if they can find that information online, not realizing the attackers could obtain the same information online when crafting phishing emails "*I can't remember the color of the bank logo. I am going to search it up [on the internet]. I am guessing this is the exact same one here [on the internet] as the one here [in the email]*" (P18–L). On the other hand, some participants explained that they would separately log into their accounts to check the information given in the email, which is a much safer option [see H13.2] "*It seems like a bit of a risk to click it, if you want to check Facebook I would instead just go directly to the Facebook website*" (P15–P). In the meantime, some participants wanted to check the information by contacting the relevant person/organization. Some choose to call the phone numbers given the email, not realizing that they could be calling phone scammers [see H13.3], and others choose to call the phone numbers found on the internet [see H13.4]. Others would physically visit the entity (e.g., bank) to verify

the email [see H13.5]. A participant even explained that she would ask a friend for a second opinion on the legitimacy of the email [see H13.6] "*It ticks all of my boxes ... But I still get the feeling that maybe the timing is a little bit off. What I will do is ask a friend*" (P15–P). We also encountered situations where participants wanted to verify the information by replying to emails, not realizing that it is unsafe to reply to a phishing email [see H13.7]. On the other hand, some participants explained that their strategy to verify doubtful emails is to click on the given links and carefully observe the redirecting process or the landing page [see H13.8] "*Whenever there's a weird thing, there's a phishing link, you click on it, and they show like redirecting. Then another website will pop up. So I'll just close the tab at that point.*" (P18–P) In summary, the following relationships can exist.

H13.1 - H13.6, H13.8: Perceived trust [by validating email content against the information on the internet, by validating the email content against the information on personal accounts, by calling phone numbers obtained from the email by calling phone numbers sourced from the internet, by physically visiting the entity, by consulting others, based on the landing pages/process] positively affects the intention to respond
H13.7: Feeling safe to validate the email by replying to the email positively affects the intention to respond

M. H14: Effects of "personal habits" on "intention to respond to an email"

Our findings reveal that user habits can affect their intention to respond negatively and positively. We observed several occasions when people make hasty decisions regarding email responses [see H14.1]. There are several reasons; some generally do not read the emails they get and believe it is easy to click on links without much thinking "*It's in general when someone shares a document with me? I click it and look*" (P06–P). Some get stressed whenever there is an unattended email notification in their inbox. On the other hand, some participants indicated they would think it is much easier to click on the link during a busy day "*If I was really busy, I would just click on the link*" (P15–P). Some explained that they would consider their priorities and the relative importance of the email if they are having a busy day [see H14.2]: "*It will depend upon my priorities. If I'm free, I'll open it. If not, I will suspend it. Or I may ignore it completely*" (P08-L).

On the other hand, some seem to completely trust in anti-virus software installed on the computer and tend to transfer the responsibility of detecting and filtering phishing emails to them [see H14.3] "*There is something [anti virus software] to take care of this*" (P14–L). We found another set of users who avoid clicking on links in the email client and use the relevant apps installed in their mobiles to take any required actions [see H14.4]. They use those mobile apps as they consider it to be the most convenient option for them: "*It's easier. The main thing I use Facebook for is Messenger. So it's just there*" (P14–L). Although these users do not use mobile apps, considering safety reasons, their habit of using mobile apps for convenience reduces the possibility of them falling prey to phishing emails.

Some participants explained that they are generally suspicious about any email that they receive in their inbox and avoid responding to any email even when they feel the email is legitimate [see H14.5]. Furthermore, some are extra cautious about specific types of emails (e.g., alert/banking emails) [see H14.6]. They feel that attackers launch their attacks usually through those types of emails; hence do not want to respond to those types of emails through the email app. Instead, they may use the mobile app or separately log in to the website to carry out any required actions.

H14.1 - H14.3: [Hasty decision-making, Relative importance given to the email, Complete trust on anti-virus software] positively affects the intention to respond
H14.4 - H14.6: [The use of the mobile app for convenience, Skepticality about any email, Extra vigilance about alert/banking emails] negatively affects the intention to respond

N. H15: Effects of "emotional attachments" on "intention to respond an email"

Our data reveal that people could respond to emails based on emotions even without considering email legitimacy or overlooking judgments they made about email legitimacy. More specifically, people could be driven to respond to emails when they are happy or excited to receive those emails [see H15.1]. For example, we observed that when participants are offered a job, they are keen to instantly click on the links to view more details or accept the offer "*I'm so happy about the offer. I will respond soon, I'll accept the offer*" (P01–P). Some emails make people curious, especially if some information is directly not visible in the email [see H15.2]. Curiosity causes people to respond to emails faster "*If I am curious to know more details, I would go and click order details and see all the details*" (P01–P). Similarly, fear of losing assets, information, and access to accounts could drive people to respond to emails immediately without a conscious decision of the action [see H15.3] "*How has somebody has added this email to my account. This worries me because if my account is linked to PayPal, all my money will be easily transferred. I would click on this link*" (P03–P).

In our data, anxiety related to work-related priorities and relationships is a very clear driver that made participants want to respond to emails even where they doubted the legitimacy [see H15.4]. For example, when the participants saw an email from their supervisor, they wanted to respond immediately "*I wouldn't question ... The supervisor has, you know, superiority. So, it feels like I have to get this resolved today as I want to impress*" (P16–P). On the other hand, we observed mixed reactions when participants believed that they received an email from their family. For example, at times, they explained the desire to respond to such emails immediately as they feel it is important to attend to family matters urgently [see H15.5]. On the other hand, some may take a more relaxed approach to respond personal emails as they feel that they can always attend to such emails later without any repercussions [see H15.6] "*It's quite informal, and I feel like I could deal with it at a later without any repercussion*" (P05–P).

In Section IV-I, we described how people could get angry or frustrated after discovering a phishing email in their inbox.

We observed that anger or frustration could lead to mixed reactions in terms of email response. In such situations, mostly people tend to delete or ignore the email [see H15.7]. "*I feel this is more like spam. I think I probably just deleted it*" (P19–P). On the other hand, we observed situations where participants get frustrated after concluding that they have received a phishing email, clicking on the email links with the intention of avoiding future phishing encounters [see H15.8]. For example, a participant explained that he wanted to click on the unsubscribe link to make sure he would not receive such phishing emails from the same sender again "*I want to click here ... I know that this is a scam email ... But I want to unsubscribe from this newsletter*" (P06–P). Therefore, based on our data, the following relationships exist:

H15.1 -15.5, 15.8: [Happiness and excitement, Curiosity, Fear of losing assets, information and access to accounts, Anxiety about maintaining work priorities and relationships, Anxiety to maintain personal relationships, Motivation to avoid future phishing emails based on anger or frustration] positively affects the intention to respond
H15.6 -15.7: [Relaxed approach to maintaining personal relationships, Motivation to avoid current email based on anger or frustration] negatively affects the intention to respond

V. Limitations

The theoretical model presented in this paper emerged from the qualitative data we collected through in-depth think-aloud sessions and interviews. We reached theoretical saturation (Figure 2) which indicates adding more participants from the same group will not influence the model. Furthermore, as explained in Section III-C, we ensured the completeness and accuracy of the model in terms of the collected data through negative case analysis. However, it is also important to note that qualitative studies are inherently interpretive, and the findings are based on the studied context, thus challenging to generalize. Furthermore, as explained in Section IV, as this study is an exploration into individuals' decision-making processes that influence their email response decisions, all concepts and relationships that emerged from the qualitative data (involving 19 participants where each participant went through 12 emails during the study) are included in the model regardless of their frequencies. Future studies could validate the model and also evaluate the significance of the elements of the user's email response decision-making process and their relationships uncovered through our data using large-scale surveys also considering diverse demographics, specific phishing-attacks, and real-life settings.

Our participants were university students; who are usually more tech-savvy and have a higher critical thinking ability [44], [81]. Previous studies also have revealed demographics like education would not influence or correlate to phishing susceptibility because an individual's knowledge does not reflect on behavior [44]. Nevertheless, similar to many other studies on phishing conducted using a sample from a university population, we cannot compare or make conclusions about other populations with different characteristics.

As explained in Section III, instead of surveys or retrospective interviewing, a role-play-based think-aloud method is better suited to achieve the aim of this research. Although a role-play scenario lacks real-world validity as users do not actually receive the emails in their usual inbox, the role-play method provided us advantages over conducting the study using real phishing emails. By the analysis of both phishing and legitimate emails, we were able to obtain much deeper insights into the underlying decision-making processes of people rather than only considering actions performed on phishing emails as in studies with 'real phishing' emails. Additionally, this setup also allowed us to employ twelve emails per participant rather than one phishing email which allowed us to present diverse types of emails. As a result, the model that emerged from our data enables us to identify general email decision-making flaws that attackers could potentially exploit to launch successful phishing attacks.

Similar to other phishing-related studies that used role-play in their study design [1], [23], [63], we designed the role-play scenario and the emails to be realistic and to match the participants as closely as possible. In fact, we used real phishing emails, real legitimate emails, and real phishing URLs for the study and adapted them to suit the given scenario. Furthermore, as mentioned in Section III, through the simulated email client, the dates and times specified in the email content were automatically adjusted based on the time participants opened those emails. Similar to previous research [53], [73], we also told participants the nature of the experiment from the start. Although the researcher clearly explained that the goal is to understand how they make email responses and not to measure how well they perform, the instructions may have primed them to examine emails more vigilantly than they usually do as they could be conscious about their thought processes being monitored by the researchers. In addition to fulfilling the ethical and responsible research of informed consent, there can be advantages to participants knowing they are taking part in a phishing-related experiment focusing on their decision-making processes compared to a deceptive study. For instance, we observed participants speaking to us effortlessly regarding their decision-making process including but not limited to how how their previous phishing email exposures and phishing education affected their email response decisions.

VI. Discussion

A deeper understanding of human decision-making, misconceptions, and user assumptions is crucial in the design of anti-phishing education, training, and awareness intervention. Through this work, we extend the state-of-the-art by providing a qualitative user study to propose a theoretical model that interprets different elements and relationships of users' email response decision-making process that influence their email responses. The theoretical model allows us to identify several general email decision-making flaws that attackers could potentially exploit to launch successful phishing attacks. Similar to previous studies [54], [81], our study highlights that people tend to respond to targetted emails with detailed information contextually aligned to their situations [25]–[27], [33], [77], [80], looking professional [26] or looking familiar [42], [77]. The role of personal characteristics and habits in email decision-making is also highlighted on several occasions [7], [27], [36], [49], [65].

A. The novelty of the findings

Apart from confirming what is already known, our study also provides novel or more profound insights into elements of the users' email response decision-making process. Moreover, the developed model not only focuses on the high-level categories (e.g., personal habits) but also on the lower concepts (e.g., trust anti-virus software, the use of mobile apps for convenience, extra vigilance about alert/banking emails) and their relationships with categories or latent variables. Hence, our model can interpret how different elements of people's email response decision-making processes could positively and negatively influence their email response behavior, which was lacking in previous literature. For example, although previous research has hinted at the importance of personal habits [34], [65], [71], [72] and emotions [1], [20], [34], [41], [82] in phishing prevention, interpretation of how different types of emotions and personal habits influence response decisions is lacking. Our results help to bridge this gap. More specifically, our results suggest certain habits (e.g., having complete trust in anti-virus software) could increase the possibility of responding to emails, and other habits (e.g., extra vigilance about alert/banking emails) could reduce this possibility (see Section IV-M). In the phishing context, this implies that certain habits reduce the possibility of falling for phishing while others increase this possibility. However, such detailed insights were missing in previous literature.

Our findings (see Section IV-L) also suggest that people may not make a final decision on email legitimacy while going through an email and may intend to validate the email before deciding how they want to respond. While some validation techniques people utilize are safe, some techniques (e.g., searching for information online, calling the phone numbers given in the email, and checking with others) could make them susceptible to potential phishing emails. Furthermore, our study reveals that people could learn correct and incorrect strategies to detect phishing emails from their past phishing encounters (see Section IV-H) and may even decide on the legitimacy of emails based on auxiliary security content available in emails (see Section IV-G).

In terms of URLs, several previous studies have investigated how people read and interpret phishing URLs without the email context [3], [6]. Even the studies conducted in the phishing email context [11], [75], do not provide evidence-based insights into users' decision-making processes related to link legitimacy determination. Our findings help to bridge this gap. For example, we found that people's decisions on email legitimacy could be based on the perceived non-mandatory nature of the links and even the appearance of buttons. Furthermore, the presence or absence of email links is also a deciding factor of email legitimacy. Our participants assumed that emails without any links could be safe not realizing that attackers could still use such emails to make people download malware or reveal sensitive information.

In terms of the email source, in line with previous research [27], [30], [45], [49], [55], [72], our findings reveal how people perceive the sender's legitimacy has a pivotal in trust they place in a received email (see Section IV-A). For example, previous research explains that people are more likely to accept a message when the source presents itself as credible. Our findings also go in line with the findings in [83] which found that people do not necessarily have a blind spot for email source details but instead do not properly recognize deception tactics commonly employed by phishing attackers. However, existing research is limited in explaining how people perceive the legitimacy of the sender by interpreting different parts of the sender's address (e.g., based on the display name of the sender address, sub-domain of the sender address, email addresses specified in the email body, etc.) and confusions they have about determining the sender's legitimacy. Our findings help to bridge this research gap.

B. Recommendations for anti-phishing education, training, and awareness intervention design

In this section, we discuss the implications of our findings to designers and researchers working in the area of anti-phishing education, training, and awareness interventions.

1) Facilitating to eliminate misconceptions and invalid assumptions: Our study findings point to several misconceptions and invalid assumptions users have with respect to strategies that phishing attackers use and their capacities. For example, we saw that people often assume that phishing emails always contain URLs; hence, the absence of email links tends to create a false sense of security in email recipients (see H6.1). Some assume that phishing attackers could never have access to detailed information about a specific situation and/or information specific to the participants, such as order details of an Amazon order (see H5.2). Some believe that anti-virus software can protect them from phishing attacks (see H14.3), and some assume that phishing attackers usually launch their attacks (see H14.6) by mainly using specific types of emails (e.g., alert/banking emails). We saw that such misconceptions and assumptions could drive people to make unsafe response decisions even when they know how to detect phishing emails correctly or have some suspicion about the received email due to other phishing cues that they noticed in the email. Therefore, anti-phishing education and training should not only focus on guiding people on identifying a phishing email correctly but also eliminate the misconceptions and invalid assumptions they have about the strategies and capacities of phishing attackers.

2) Tailored anti-phishing education, training and awareness: Our results provide insights into the diversity and complexity of how people make email responses. For example, our results point to situations where some people struggle to identify the legitimacy of emails, some struggle to validate the emails, and some struggle to take safe actions even after making correct legitimacy judgments. Hence, it's important to focus on targeted training and awareness without a one-size-fits-all approach. We see the value of using personas [52], in an organization setting, for designing tailored education, training, and awareness interventions. Personas are fictitious representations of user groups, their goals, and preferences for bridging the gap between designers and the end-users they are designing for [58]. While personas have been lauded for their benefits, they are rarely used in the context of phishing prevention. The development of personas requires a deeper understanding of the behavioral and demographic characteristics of the users. Hence, we anticipate that the knowledge generated from this study, along with other work that looks into the demographic and personality of people who fall for phishing emails [23],

[36], [39] could serve as a starting point for designing effective tailored anti-phishing interventions in the future.

3) Shifting focus from accurate email legitimacy judgements to secure email responses: Our theoretical model suggests that perceived email legitimacy may not be the sole influencer of email response decisions. However, frequently anti-phishing education, training, and awareness interventions [59], [78], [79] often focus only on people's email legitimacy judgments in their study designs and/or in their evaluations. Our results indicate other important factors that should be considered independently or in conjunction with people's email legitimacy judgments to ensure safer email response decisions. For example, our results indicate people could fall prey to phishing even after identifying phishing emails correctly (see H15.8 in Section IV-N). Therefore, in the future, we expect the focus of the design and evaluation of both research and development on anti-phishing education, training, and awareness interventions to shift from *accurate email legitimacy judgments* to *secure email responses*.

4) Facilitating safe email validation: Given that the trust perceived based on email validation techniques could drive people's email response behavior (see Section IV-L), it is important to make sure people validate emails in safer ways. Unfortunately, our study provides evidence that people use unsafe techniques to validate emails when they have doubts about email legitimacy. To the best of our knowledge, this information was missing in previous literature and is an important part of the puzzle that explains why people fall prey to phishing emails. Therefore, we see value in tools being developed to facilitate employees in organizational settings to validate emails, specifically when they are doubtful about email legitimacy. Jenkins et al. [35] proposed initial UI designs for a tool to support people who report phishing so that they can confidentially take appropriate action. Althobaiti et al. [6] designed a usable report based on the information professionals use to support users in deciding if potential phishing URLs are or are not safe to click on. We anticipate such tools to be integrated into email clients and adapted to provide contextualized just-in-time advice to the users on all aspects of an email that they have doubts. For example, our study provides evidence that users struggle in interpreting the reply-to address and sender address. Users can also be unaware of how to validate the organizational logos visible in emails correctly and, as a result, trust outdated logos based on the results of internet searches they perform. Furthermore, anti-phishing education and training material should also be updated to teach users about safe and unsafe email validation techniques specifically.

5) Giving more prominence to diverse personal habits and emotions in tool design: Our results provide insights into how different emotions (see Section IV-N) and personal habits (see Section IV-M) can positively and negatively influence people's response behaviors. For example, certain emotions such as happiness and excitement, anxiety about maintaining work priorities and relationships, and anxiety about maintaining personal relationships positively affect the intention to respond to phishing emails. Emotions can also override people's legitimacy judgments and drive them to make undesirable responses to phishing emails (see H15.8). As a result, attackers could craft emails to manipulate users' emotions in order to increase the possibility of them falling for those emails. However, despite the importance, previous research work has paid less attention to personal habits and emotions in the design of anti-phishing training, education, and awareness tools [5], [25]. We see value in providing people opportunities to self-assess their habits and emotions and the impact of those on their email response decisions. The knowledge generated by such self-assessment tools can be used to create self-awareness, facilitating people to be more vigilant about their response behaviors and also providing targeted anti-phishing training. Furthermore, the knowledge arising from this research, together with advancements in automatic sentiment and emotion recognition in the email context [31] could be used to guide users to make safer email responses in the wild.

6) Facilitating to assess the validity of self-learned strategies: Our findings in Section IV-L reveal several inaccurate strategies that people learned through past phishing encounters that they use to detect phishing emails. They could apply those strategies to new emails leading to unsafe email responses to phishing emails (see H8.2). This highlights the need for providing support for users to self-validate strategies they have learned to prevent them from applying those strategies in the future. We see value in providing tool support to users to validate self-learned phishing detection strategies. For example, we anticipate advances in natural language processing, such as semantic similarity [13], can be utilized to develop intelligent systems to process user-defined phishing email detection strategies and provide advice or feedback on whether those strategies are safe. Furthermore, role-playing simulation games as in [79] could be expanded to focus on correcting the self-learned incorrect strategies revealed through this study.

VII. Conclusion

In this paper, we investigate in-depth how people make email response decisions while reading their emails. Analysis of the collected qualitative data enabled us to develop a theoretical model that describes how people can be driven to respond to emails by clicking on email links and replying to or downloading attachments based on people's email response decision-making elements and their relationships. Based on an improved understanding of how people make email responses, this study enables us to identify how people can be susceptible to manipulation, even in our controlled experiment environment. We proposed five concrete enhancements to state-of-the-art anti-phishing education, training, and awareness tools to support users in making safe email responses. Among others, we suggest that the goal of anti-phishing education, training, and awareness tools should shift from accurate email legitimacy judgments to secure email responses. Therefore, we believe our work lays the foundation for improving future anti-phishing interventions to make a significant difference in how we prevent phishing email attacks in the future.

Acknowledgment

This work was carried out while the first author was working at the Centre for Research on Engineering Software Technologies (CREST), University of Adelaide.

REFERENCES

[1] H. Abroshan, J. Devos, G. Poels, and E. Laermans, "Covid-19 and phishing: effects of human emotions, behavior, and demographics on the success of phishing attempts during the pandemic," *IEEE Access*, vol. 9, pp. 121 916–121 929, 2021.

[2] M. Ackerley, B. Morrison, K. Ingrey, M. Wiggins, P. Bayl-Smith, N. Morrison *et al.*, "Errors, irregularities, and misdirection: Cue utilisation and cognitive reflection in the diagnosis of phishing emails," *Australasian Journal of Information Systems*, vol. 26, 2022.

[3] S. Albakry, K. Vaniea, and M. Wolters, "What is this url's destination? empirical evaluation of web users' url reading," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, 12 2019.

[4] L. Allodi, T. Chotza, E. Panina, and N. Zannone, "The need for new antiphishing measures against spear-phishing attacks," *IEEE Security & Privacy*, vol. 18, no. 2, pp. 23–34, 2019.

[5] M. Alshaikh, S. B. Maynard, and A. Ahmad, "Applying social marketing to evaluate current security education training and awareness programs in organisations," *Computers & Security*, vol. 100, p. 102090, 2021.

[6] K. Althobaiti, N. Meng, and K. Vaniea, "I don't need an expert! making url phishing features human comprehensible," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–17.

[7] E. Ayaburi and F. K. Andoh-Baidoo, "Understanding phishing susceptibility: an integrated model of cue-utilization and habits," 2019.

[8] Berkeley University of California. (2022) berkeley phishing examples archive. [Online]. Available: https://security.berkeley.edu/education-awareness/phishing/phishing-examples-archive

[9] I. Bose and A. C. M. Leung, "Do phishing alerts impact global corporations? a firm value analysis," *Decision Support Systems*, vol. 64, pp. 67–78, 2014.

[10] A. E. Brodsky, "Negative case analysis," *The SAGE encyclopedia of qualitative research methods*, vol. 2, p. 552, 2008.

[11] M. Butavicius, R. Taib, and S. J. Han, "Why people keep falling for phishing scams: The effects of time pressure and deception cues on the detection of phishing emails," *Computers & Security*, vol. 123, p. 102937, 2022.

[12] A. Caspi, M. Sayag, M. Gross, Z. Weinstein, and S. Etgar, "The effects of personal values and message values on vulnerability to phishing," *Personality and Individual Differences*, vol. 186, p. 111335, 2022.

[13] D. Chandrasekaran and V. Mago, "Evolution of semantic similarity—a survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 2, pp. 1–37, 2021.

[14] R. Chen, J. Gaia, and H. R. Rao, "An examination of the effect of recent phishing encounters on phishing susceptibility," *Decision Support Systems*, vol. 133, p. 113287, 2020.

[15] CISCO, "2021 cyber security threat trends- phishing, crypto top the list," CISCO, Tech. Rep., 2021.

[16] A. Compitition and C. Commission. (2022) Identity theft. [Online]. Available: https://www.scamwatch.gov.au/types-of-scams/attempts-to-gain-your-personal-information/identity-theft

[17] X. Cui, Y. Ge, W. Qu, and K. Zhang, "Effects of recipient information and urgency cues on phishing detection," in *International Conference on Human-Computer Interaction*. Springer, 2020, pp. 520–525.

[18] G. Desolda, L. S. Ferro, A. Marrella, T. Catarci, and M. F. Costabile, "Human factors in phishing attacks: A systematic literature review," *ACM Computing Surveys (CSUR)*, vol. 54, no. 8, pp. 1–35, 2021.

[19] M. Fernando and N. A. G. Arachchilage, "Why johnny can't rely on anti-phishing educational interventions to protect himself against contemporary phishing attacks?" *arXiv preprint arXiv:2004.13262*, 2020.

[20] A. Ferreira and S. Teles, "Persuasion: How phishing emails can influence users and bypass security measures," *International Journal of Human-Computer Studies*, vol. 125, pp. 19–31, 2019.

[21] A. Franz, V. Zimmermann, G. Albrecht, K. Hartwig, C. Reuter, A. Benlian, and J. Vogt, "{SoK}: Still plenty of phish in the sea—a taxonomy of {User-Oriented} phishing interventions and avenues for future research," in *Seventeenth Symposium on Usable Privacy and Security (SOUPS 2021)*, 2021, pp. 339–358.

[22] S. Garera, N. Provos, M. Chew, and A. D. Rubin, "A framework for detection and measurement of phishing attacks," in *Proceedings of the 2007 ACM workshop on Recurring malcode*. ACM, 2007, pp. 1–8.

[23] Y. Ge, L. Lu, X. Cui, Z. Chen, and W. Qu, "How personal characteristics impact phishing susceptibility: The mediating role of mail processing," *Applied Ergonomics*, vol. 97, p. 103526, 2021.

[24] B. G. Glaser and A. L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine Transaction, 1967.

[25] S. Goel, K. Williams, and E. Dincelli, "Got phished? internet security and human vulnerability," *Journal of the Association for Information Systems*, vol. 18, no. 1, p. 2, 2017.

[26] K. K. Greene, M. P. Steves, M. F. Theofanos, and J. Kostick, "User context: an explanatory variable in phishing susceptibility," in *in Proc. 2018 Workshop Usable Security*, 2018.

[27] F. L. Greitzer, W. Li, K. B. Laskey, J. Lee, and J. Purl, "Experimental investigation of technical and human factors related to phishing susceptibility," *ACM Transactions on Social Computing*, vol. 4, no. 2, pp. 1–48, 2021.

[28] A.-P. W. Group, "Phishing activity trends report: 1st quarter 2020," Anti-Phishing Working Group, Tech. Rep., 2020.

[29] B. B. Gupta, N. A. Arachchilage, and K. E. Psannis, "Defending against phishing attacks: taxonomy of methods, current issues and future directions," *Telecommunication Systems*, vol. 67, no. 2, pp. 247–267, 2018.

[30] Z. M. Hakim, N. C. Ebner, D. S. Oliveira, S. J. Getz, B. E. Levin, T. Lin, K. Lloyd, V. T. Lai, M. D. Grilli, and R. C. Wilson, "The phishing email suspicion test (pest) a lab-based task for evaluating the cognitive mechanisms of phishing detection," *Behavior research methods*, vol. 53, no. 3, pp. 1342–1352, 2021.

[31] Z. Halim, M. Waqar, and M. Tahir, "A machine learning-based investigation utilizing the in-text features for the identification of dominant emotion in an email," *Knowledge-based systems*, vol. 208, p. 106443, 2020.

[32] B. Harrison, E. Svetieva, and A. Vishwanath, "Individual processing of phishing emails: How attention and elaboration protect against phishing," *Online Information Review*, 2016.

[33] L. Jaeger and A. Eckhardt, "Eyes wide open: The role of situational information security awareness for security-related behaviour," *Information Systems Journal*, vol. 31, no. 3, pp. 429–472, 2021.

[34] A. Jayatilaka, N. A. G. Arachchilage, and M. A. Babar, "Falling for phishing: An empirical investigation into people's email response behaviors," in *International Conference on Information Systems (ICIS) 2021 Proceedings*, 2021.

[35] A. Jenkins, N. Kokciyan, and K. E. Vaniea, "Phished: Automated contextual feedback for reported phishing," in *18th Symposium on Usable Privacy and Security*. Usenix, 2022.

[36] H. S. Jones, J. N. Towse, N. Race, and T. Harrison, "Email fraud: The search for psychological predictors of susceptibility," *PloS one*, vol. 14, no. 1, p. e0209684, 2019.

[37] I. Kirlappos and M. A. Sasse, "Security education against phishing: A modest proposal for a major rethink," *IEEE Security & Privacy*, vol. 10, no. 2, pp. 24–32, 2011.

[38] Y. Kwak, S. Lee, A. Damiano, and A. Vishwanath, "Why do users not report spear phishing emails?" *Telematics and Informatics*, vol. 48, p. 101343, 2020.

[39] P. Lawson, C. J. Pearson, A. Crowson, and C. B. Mayhorn, "Email phishing and signal detection: How persuasion principles and personality influence response patterns and accuracy," *Applied ergonomics*, vol. 86, p. 103084, 2020.

[40] P. Lawson, O. Zielinska, C. Pearson, and C. B. Mayhorn, "Interaction of personality and persuasion tactics in email phishing attacks," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 61, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2017, pp. 1331–1333.

[41] N. LeFranc and A. Savoli, "Factors influencing employees' susceptibility to phishing emails: The role of emotions," in *Proc. 13th Medit. Conf. Inf. Syst.(MCIS)*, 2019, pp. 1–8.

[42] J. Lim, L. Zhou, and D. Zhang, "Verbal deception cue training for the

detection of phishing emails," in *2021 IEEE International Conference on Intelligence and Security Informatics (ISI)*. IEEE, 2021, pp. 1–3.

[43] T. Lin, D. E. Capecci, D. M. Ellis, H. A. Rocha, S. Dommaraju, D. S. Oliveira, and N. C. Ebner, "Susceptibility to spear-phishing emails: Effects of internet user demographics and email content," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 26, no. 5, pp. 1–28, 2019.

[44] Z. Liu, L. Zhou, and D. Zhang, "Effects of demographic factors on phishing victimization in the workplace." in *PACIS*, 2020, p. 75.

[45] X. R. Luo, W. Zhang, S. Burd, and A. Seazzu, "Investigating phishing victimization with the heuristic–systematic model: A theoretical framework and an exploration," *Computers & Security*, vol. 38, pp. 28–38, 2013.

[46] A. Mai, K. Pfeffer, M. Gusenbauer, E. Weippl, and K. Krombholz, "User mental models of cryptocurrency systems-a grounded theory approach," in *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, 2020, pp. 341–358.

[47] G. Misra, N. A. G. Arachchilage, and S. Berkovsky, "Phish phinder: A game design approach to enhance user confidence in mitigating phishing attacks," *arXiv preprint arXiv:1710.06064*, 2017.

[48] K. A. Molinaro and M. L. Bolton, "Evaluating the applicability of the double system lens model to the analysis of phishing email judgments," *computers & security*, vol. 77, pp. 128–137, 2018.

[49] G. D. Moody, D. F. Galletta, and B. K. Dunn, "Which phish get caught? an exploratory study of individuals' susceptibility to phishing," *European Journal of Information Systems*, vol. 26, no. 6, pp. 564–584, 2017.

[50] M. M. Moreno-Fernández, F. Blanco, P. Garaizar, and H. Matute, "Fishing for phishers. improving internet users' sensitivity to visual deception cues to prevent electronic fraud," *Computers in Human Behavior*, vol. 69, pp. 421–436, 2017.

[51] G. Nasser, B. W. Morrison, P. Bayl-Smith, R. Taib, M. Gayed, and M. W. Wiggins, "The role of cue utilization and cognitive load in the recognition of phishing emails," *Frontiers in big data*, vol. 3, p. 546860, 2020.

[52] T. Neate, A. Bourazeri, A. Roper, S. Stumpf, and S. Wilson, "Co-created personas: Engaging and empowering users with diverse needs within the design process," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–12.

[53] J. Nicholson, L. Coventry, and P. Briggs, "Can we fight social engineering attacks by social means? assessing social salience as a means to improve phish detection," in *Thirteenth Symposium on Usable Privacy and Security ({SOUPS} 2017)*, 2017, pp. 285–298.

[54] K. Parsons, A. McCormac, M. Pattinson, M. Butavicius, and C. Jerram, "Phishing for the truth: A scenario-based experiment of users' behavioural response to emails," in *IFIP International Information Security Conference*. Springer, 2013, pp. 366–378.

[55] K. Pfeffel, P. Ulsamer, and N. H. Müller, "Where the user does look when reading phishing mails–an eye-tracking study," in *International Conference on Human-Computer Interaction*. Springer, 2019, pp. 277–287.

[56] PhishTank. (2022) Phishtank. [Online]. Available: https://phishtank.org/

[57] N. Pilavakis, A. Jenkins, N. Kökciyan, and K. Vaniea, ""i didn't click": What users say when reporting phishing," in *Symposium on Usable Security and Privacy (USEC) 2023*. The Internet Society, 2023, pp. 1–13.

[58] R. M. Quintana, S. R. Haley, A. Levick, C. Holman, B. Hayward, and M. Wojan, "The persona party: using personas to design for learning at scale," in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2017, pp. 933–941.

[59] B. Reinheimer, L. Aldag, P. Mayer, M. Mossano, R. Duezguen, B. Lofthouse, T. Von Landesberger, and M. Volkamer, "An investigation of phishing awareness and education over time: When and how to best remind users," in *Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020)*, 2020, pp. 259–284.

[60] S. Salloum, T. Gaber, S. Vadera, and K. Shaalan, "A systematic literature review on phishing email detection using natural language processing techniques," *IEEE Access*, vol. 10, pp. 65 703–65 727, 2022.

[61] D. M. Sarno and M. B. Neider, "So many phish, so little time: Exploring email task factors and phishing susceptibility," *Human Factors*, p. 0018720821999174, 2021.

[62] Scam Detector. (2022) Scam detector. [Online]. Available: https://www.scam-detector.com/

[63] K. Schiller, F. Adamsky, and Z. Benenson, "Towards an empirical study to determine the effectiveness of support systems against e-mail phishing attacks," in *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 2023, pp. 1–15.

[64] Sensors Tech Forum. (2022) Sensors tech forum. [Online]. Available: https://sensorstechforum.com/

[65] H. Shahbaznezhad, F. Kolini, and M. Rashidirad, "Employees' behavior in phishing attacks: what individual, organizational, and technological factors matter?" *Journal of Computer Information Systems*, vol. 61, no. 6, pp. 539–550, 2021.

[66] S. Sheng, M. Holbrook, P. Kumaraguru, L. F. Cranor, and J. Downs, "Who falls for phish?: a demographic analysis of phishing susceptibility and effectiveness of interventions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 373–382.

[67] A. A. Smith-Ditizio and A. D. Smith, "Computer fraud challenges and its legal implications," in *Advanced Methodologies and Technologies in System Security, Information Privacy, and Forensics*. IGI Global, 2019, pp. 152–165.

[68] T. Stojnic, D. Vatsalan, and N. A. Arachchilage, "Phishing email strategies: Understanding cybercriminals' strategies of crafting phishing emails," *Security and Privacy*, vol. 4, no. 5, p. e165, 2021.

[69] A. L. Strauss and J. M. Corbin, *Basics of Qualitative Research : Techniques and Procedures for Developing Grounded Theory*, 2nd ed. Sage, 1998.

[70] Tessian, "Phishing statistics report 2021 phishing statistics report 2020," Tessian, Tech. Rep., 2021.

[71] A. Vishwanath, B. Harrison, and Y. J. Ng, "Suspicion, cognition, and automaticity model of phishing susceptibility," *Communication Research*, vol. 45, no. 8, pp. 1146–1166, 2018.

[72] A. Vishwanath, T. Herath, R. Chen, J. Wang, and H. R. Rao, "Why do people get phished? testing individual differences in phishing vulnerability within an integrated, information processing model," *Decision Support Systems*, vol. 51, no. 3, pp. 576–586, 2011.

[73] J. Wang, Y. Li, and H. R. Rao, "Overconfidence in phishing email detection," *Journal of the Association for Information Systems*, vol. 17, no. 11, p. 1, 2016.

[74] ——, "Coping responses in phishing detection: an investigation of antecedents and consequences," *Information Systems Research*, vol. 28, no. 2, pp. 378–396, 2017.

[75] R. Wash, "How experts detect phishing scam emails," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. CSCW2, pp. 1–28, 2020.

[76] R. Wash and M. M. Cooper, "Who provides phishing training?: Facts, stories, and people like me," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. 492.

[77] R. Wash, N. Nthala, and E. Rader, "Knowledge and capabilities that non-expert users bring to phishing detection," in *Symposium on Usable Privacy and Security*, 2021.

[78] B. W. Weaver, A. M. Braly, and D. M. Lane, "Training users to identify phishing emails," *Journal of Educational Computing Research*, vol. 59, no. 6, pp. 1169–1183, 2021.

[79] Z. A. Wen, Z. Lin, R. Chen, and E. Andersen, "What. hack: engaging anti-phishing training through a role-playing phishing simulation game," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[80] E. J. Williams, J. Hinds, and A. N. Joinson, "Exploring susceptibility to phishing in the workplace," *International Journal of Human-Computer Studies*, vol. 120, pp. 1–13, 2018.

[81] E. J. Williams and D. Polage, "How persuasive is phishing email? the role of authentic design, influence and current events in email judgements," *Behaviour & Information Technology*, vol. 38, no. 2, pp. 184–197, 2019.

[82] M. Workman, "Wisecrackers: A theory-grounded investigation of phishing and pretext social engineering threats to information security,"

Journal of the American society for information science and technology, vol. 59, no. 4, pp. 662–674, 2008.

[83] S. Zheng and I. Becker, "Presenting suspicious details in {User-Facing} e-mail headers does not improve phishing detection," in *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*, 2022, pp. 253–271.