

Too Good to Ignore: How Directional Goals Shape Scam Perception

Diya Krishnan, Zoie Sloneker, Yun Chieh Yang
Carnegie Mellon University

1 Introduction

Online scams remain a persistent threat across a wide range of digital communication environments [2]. Fraudulent job offers, phishing emails, impersonation messages, and other deceptive communications increasingly resemble legitimate interactions, making them difficult to identify at a glance [3]. Although public awareness campaigns and security advice have improved familiarity with common scam tactics, individuals continue to engage with suspicious messages at high rates [1].

One possible explanation is that scam perception is shaped not only by knowledge or vigilance, but also by motivation. When a message aligns with something a person wants to be true, they may interpret ambiguity in a favorable way, ignoring warning signs. This pattern is consistent with motivated reasoning, in which people selectively interpret information in ways that support desired conclusions. The interaction between personal goals and visible warning signs may help explain why even the most cautious individuals can sometimes engage with fraudulent communications.

Understanding the connection between goal-driven reasoning and susceptibility to scam communication can help create new training materials to help users avoid falling for scams better. We would like to help users avoid falling for the trap of reasoning using only a biased subset of knowledge and instead be more objective in their reasoning, regardless of any desirable outcomes. The ultimate goal of this project is to provide effective guidance on red flags and warning signs.

This paper examines how motivated reasoning shapes scam perception by exploring how people evaluate and respond to suspicious communications that align with their personal goals or desires, with particular attention to the roles of directional goal strength and red flag salience. We conducted semi-structured in-person virtual interview sessions through Zoom and Google Meet with 12 participants to analyze how they arrived at their judgments regarding the messages. We address the following research questions.

RQ1: Does directional goal strength (how much a participant

wants a scenario to be real) influence their willingness to engage with a potentially fraudulent message?

RQ2: Does red-flag salience moderate the effect of directional goal strength on participants' engagement decisions?

2 Methods

We conducted a within-subjects semi-structured interview study approved by the CMU IRB (STUDY2026_00000142). Eighteen adults were screened for eligibility; one was excluded prior to interviewing for failing to meet inclusion criteria, and five transcripts were subsequently excluded due to insufficient English proficiency or failure to satisfy residency requirements, yielding a final analytic sample of 12 participants. Eligible participants were adults aged 18 or older residing in the United States who reported regular use of both email and smartphone messaging applications. Participants were recruited via physical flyers posted on the Carnegie Mellon University campus and through online community forums, including CMU Discord servers and LinkedIn, and received a \$10 Amazon gift card as compensation.

The study employed a 2 (Directional Goal Strength: Low vs. High) $\times$ 2 (Red Flag Salience: Ambiguous vs. Clear) within-subjects design, producing four experimental conditions. Each participant evaluated six fictional but realistic email scenarios presented in randomized order: four corresponding to each of the experimental conditions and two benign fillers included to reduce suspicion. All stimuli depicted job recruitment communications. Directional goal strength was manipulated through scenario framing: high-goal stimuli described a role plausibly desirable to the participant (e.g., a position at a well-known technology company that would likely be desirable to job-seeking young adults), while low-goal stimuli described a role with no apparent personal relevance. Red flag salience was manipulated through message content: clear-condition emails contained overt fraud indicators such as spoofed sender domains (e.g., `google.career.com` rather than `google.com`),

whereas ambiguous-condition emails embedded more subtle cues such as a mildly informal tone or minor sender inconsistencies. Sessions were conducted remotely via Zoom or Google Meet and lasted approximately 30 minutes. Participants were prompted to think aloud while evaluating each message and to indicate whether they would engage with it.

Transcripts were analyzed using inductive qualitative coding; each interview was coded independently by at least two researchers, with disagreements resolved through discussion with a third. For quantitative analysis, we report frequency counts and 2×2 contingency tables, as the sample size precluded formal statistical inference.

3 Results

Of the 12 participants, seven identified as male and five as female; seven were current CMU students and five were employed professionals, with ten reporting a confirmed STEM background.

Qualitative coding of think-aloud responses identified nine recurring themes across participant evaluations of the six stimuli (Figure 1). The most prevalent theme was suspicion about the overall job description, raised by 83% of participants (10/12), followed by urgency cues (75%, 9/12) and suspicious writing style (66%, 8/12). Several participants said they would engage after verifying the sender; P5 stated they would “search up this person first on LinkedIn” before replying. Notably, 58% of participants flagged the mode of communication itself—particularly requests to contact recruiters via WhatsApp—as an indicator of fraud.

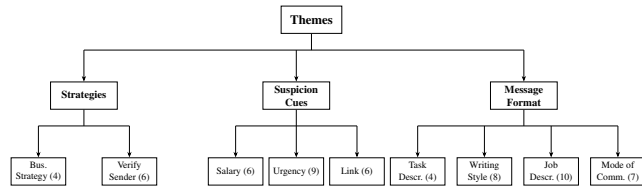


Figure 1: Tree Diagram of qualitative themes, which are divided into labels and categories. Each category has a number determining the number of people who brought it up.

Each of the participants made remarks regarding these qualitative themes. For example, P16 remarked that “normally [recruiters] don’t mention WhatsApp” in the messages they normally send to potential applicants, expressing suspicion about the communication method. P5 said that they would “search up...this person first on LinkedIn” before replying. P2 said that one of the emails looked fake because its “rates are very low.” P6 took issue with the benefits of one of the jobs, calling it “too good to be true.” P12 said that “they kind of ask you to respond urgently” when discussing the red flags of one of the emails telling users that many people are applying

for a given position. Participant P18 noticed that a link had HTTP in the front and concluded that “website is not secure.” P2 remarked that in real life, “Microsoft doesn’t reach out to people for a contract job”, contrary to what an email stated. P2 stated that one of the emails had “corporate jargon” to make it seem more trustworthy.

Quantitative counts are shown in Table 1. Across 72 evaluation opportunities, participants indicated willingness to engage on 31 occasions. Contrary to the intuition that participants in high-directional-goal conditions will be more likely to indicate that they would engage with a suspicious message than those in low-directional-goal conditions, engagement was higher in low directional goal strength conditions (21/36) than in high (10/36). Engagement rates were comparable across scam and genuine messages (20 vs. 11 engagements respectively), with 9 of the 11 engagements with genuine messages occurring in the low goal condition.

	Engage = 0	Engage = 1
Dir. Goal Str. = Low	15	21
Dir. Goal Str. = High	26	10

Table 1: Engagement counts by directional goal strength ($n = 72$ trials).

4 Limitations and Future Work

Participants engaged more with messages in low-directional-goal-strength conditions (21/36) than high (10/36). We suspect this reflects our sample’s technical background: P12 felt that a pay of 25-30 dollars for 2-3 hours of work is “too good to be true.” The small, CMU-heavy sample limits generalizability. Furthermore, due to the small sample size, we were not able to perform a complete quantitative analysis. In addition, we never asked participants how much they wanted the messages to be real; we just assumed that they wanted it the way we designed it. As a result, we had to make assumptions such that participants would be more drawn towards jobs with high salaries. Additionally, the high-goal stimuli featured roles at prominent technology companies, which may have been less personally relevant to participants without a STEM background, potentially weakening the directional goal manipulation for those individuals.

In our follow-up study we will use a more representative sample with at least 30 participants. We would also do a logistic regression analysis that would better isolate the effects of directional goal strength and red flag salience. Future protocols will use Qualtrics for participant screening (to filter non-US residents by IP address) and reorder think-aloud prompts so participants report engagement decisions before justifying them, reducing order effects. Future protocols will use a screening to find out from the participants how much they wanted the message to be real.

5 Acknowledgments

We would like to thank our mentors Alexandra Li, Dr. Lorrie Cranor, and Dr. Yuvraj Agarwal for providing us with valuable advice for our project.

References

- [1] Elif Kubilay, Eva Raiber, Lisa Spantig, Jana Cahlíková, and Lucy Kaaria. Can you spot a scam? measuring and improving scam identification ability. *Journal of Development Economics*, 165:103147, 2023.
- [2] Wai Yie Leong, Yuan Zhi Leong, and Wai San Leong. Strategies for identifying online scams. In *2024 Asian Conference on Communication and Networks (ASIAN-ComNet)*, pages 1–6, 2024.
- [3] Sinayobye Janvier Omar, Kiwanuka Fred, and Kaawaase Kyanda Swaib. A state-of-the-art review of machine learning techniques for fraud detection research. In *Proceedings of the 2018 International Conference on Software Engineering in Africa, SEiA '18*, page 11–19, New York, NY, USA, 2018. Association for Computing Machinery.

Appendix

Interview Script

Section A: Intro, consent section (1-2 min)

Thank you so much for taking time to speak with us today. My name is [Name], and joining me today is [Note taker's Name], who will be taking some notes so I can focus on our conversation. We are both students at Carnegie Mellon University.

I will now read out loud our consent form. If you agree to the consent form, please say "I agree." If you don't agree, we will stop the interview.

Consent:

[Interviewer reads aloud:] "We are inviting you to take part in a research study conducted by Diya Krishnan, Yun-Chieh Yang, and Zoie Sloneker at Carnegie Mellon University (STUDY2026_00000142). This brief (30 minute) interview will help us learn more about how adults interact with online messages.

You have been asked to take part because you are an adult aged 18 or above who frequently uses email, and online messaging apps. You can skip questions that you do not want to answer or stop at any time. Your participation is entirely voluntary.

The interview will be audio and video recorded to allow us to accurately record your responses. You are welcome to keep your camera off if you feel more comfortable that way. At the beginning of the recording, we will ask you to confirm that you agree to be recorded. There may be a risk of a breach of confidentiality and although we cannot completely remove the possibility, we have processes in place to prevent your information from being accessed by anyone not involved in the project. The audio/video recordings will be destroyed promptly following transcription of the interview and will never be shared with others or used beyond this project. We may share or use the text transcript of your interview for future research or commercial purposes; however, we will take steps to remove any information that may identify you before the transcript is shared. When we write-up our findings from the study, we will replace your name with a number and will take care to change details that could indirectly identify you or others. We will need to keep your email address so that we can send you compensation once the interview is completed.

If you have any questions, you are encouraged to ask them now. If you have questions later, please contact zsloneke@andrew.cmu.edu. If you have questions or concerns about your rights as a research participant, you can contact the Carnegie Mellon IRB at (412) 268-4721, IRB-review@andrew.cmu.edu. You will be provided with a copy of this informational script for your records. Do you consent?"

[Wait for consent. If they agree, tell them that we need to record their consent. Otherwise, we will end the interview]

You have consented. I will now begin recording and I have to confirm your agreement.

Are you in a quiet, private location so that we can avoid accidentally recording anyone else nearby?

Would it be OK if I start recording now?

[Begin Recording]

Do you consent?

"Do you have any questions before we begin our interview?"

Section B: Scam Experience, tech experience (2-3 min)

[Interviewer try to build rapport and understand the background of the participant]

Can you tell me a little about how you typically use email or messaging apps on a day-to-day basis?

Follow-up: How many e-mails or text messages do you receive in a day?

Follow-up: Do you feel like you stay on top of your messages, or do they pile up?

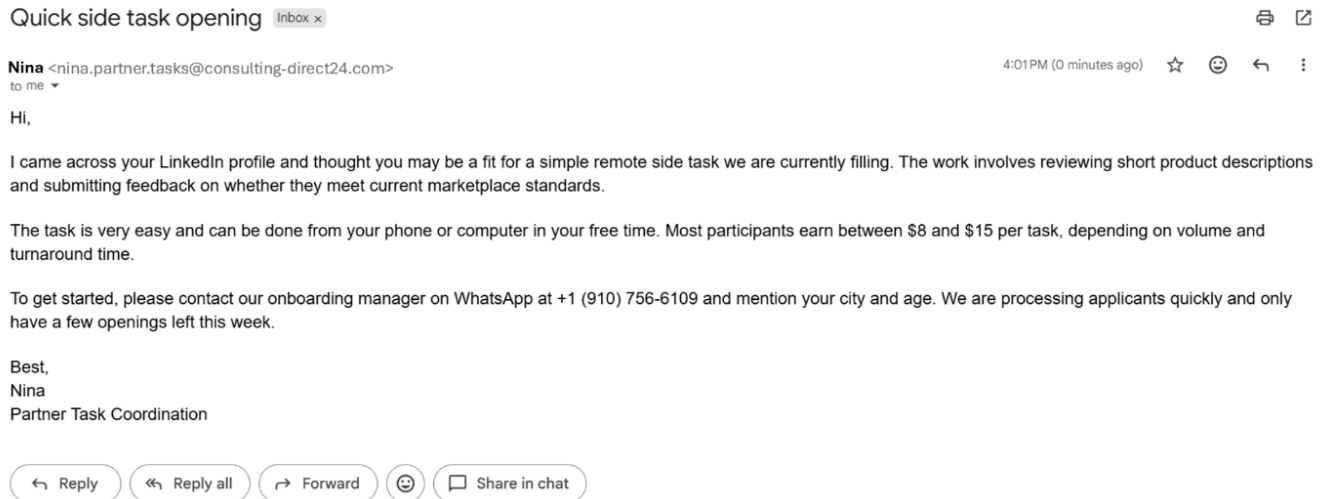
Do you receive more promotional emails, personal emails, or order updates?

Section C: Prior context and think aloud (5-6 min)

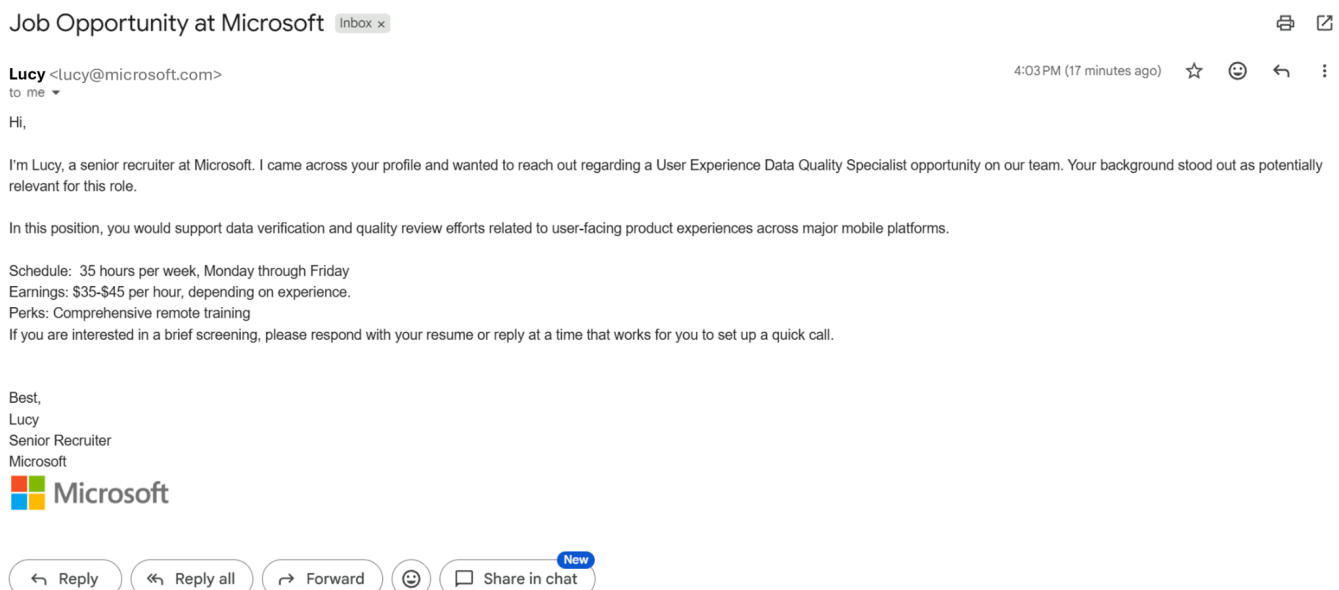
[Interviewer explains how the next few questions will flow] “I will now show you a few images of fictional emails. Let me know what you see and think about the message.”

For all questions below, if the participant believes the message to be a scam, ask them what they will do next (report the message, ignore the message, verify the message, etc). If they neither label the email as spam or not spam, ask the follow up question of “if they would follow through and respond to the message” in whatever way it requests. For each question they label as spam, ask the follow up question “Is there a part of the message you want to believe is real?”.

[Interviewer reads aloud] “Please take a look at this email [shows email]. Assume you have a LinkedIn profile, and it is marked as open to work.”



[Interviewer reads aloud] “Please take a look at this email. Assume you have accounts on platforms like LinkedIn where you are marked as looking for a job”



[Interviewer reads aloud] “Please take a look at this email [email]. Assume you have LinkedIn profiles. Assume that the senior HR coordinator of the company you work at is named John Smith.”

Leadership Development Program Inbox x



John Smith <john_smith_hr@gmail.com>

4:04 PM (31 minutes ago) ☆ 😊 ↶ ⋮

to me ▾
Hi,

Following an internal review, you have been identified as a strong candidate for our upcoming Leadership Development program. This opportunity is designed to prepare you for a managerial role in the future.

To secure your placement, please complete the following pre-enrollment verification form: <http://leadership-program-sign-up.com>

Please note that there are many candidates for this position, so we encourage you to respond promptly if interested. If you have any questions please reply to this email.

Congratulations on being recognized for your hard work!
John Smith

Senior HR Coordinator
Talent Development Division

↶ Reply ↶ Reply all ↷ Forward 😊 🗨 Share in chat New

[Interviewer reads aloud] “Please take a look at this email. Assume you have a LinkedIn profile, and it is open for work. Assume it says your name instead of [name].”

Job Recruitment Inbox x



Emily Becker <emilybecker@digital-solutions-dlt.com>

4:06 PM (34 minutes ago) ☆ 😊 ↶ ⋮

to me ▾

Hello [name]

I'm Emily, a recruitment specialist at Digital Solutions Ltd (DSL).

We recently reviewed your LinkedIn profile and we were impressed by your experience. We are currently offering a flexible, remote opportunity assisting merchants with high-quality assurance and feedback management for the TEMU marketplace.

This is a straightforward task that can be completed in your spare time. The hourly compensation ranges from \$10 to \$15, with payments processed upon project milestone completion.

If you are interested in joining our upcoming training cohort, please reach out to me.

Best,
Emily Becker
Senior Recruitment specialist
Digital Solutions Ltd

↶ Reply ↶ Reply all ↷ Forward 😊 🗨 Share in chat New

[Interviewer reads aloud] “Please take a look at this email. Assume you have accounts on platforms like LinkedIn where you are marked as looking for a job.”

Opportunity: Customer Support Engineer Co-op at Google Inbox x



Lucas Johnson <lucasjohnson@google.com>
to me ▾

4:07 PM (40 minutes ago) ☆ 😊 ↶ ⋮

Hi,

I hope you're doing well! I'm Lucas, a Talent Acquisition Sourcer at Google, and we're currently hiring for a customer co-op for Fall 2026. I came across your profile and thought you would be a great fit for this opportunity.

As a Customer Co-op at Google you will:

Work with clients to understand their AI and data needs
Design AI solutions and walk clients through how they work
Partner with engineers, product teams, and sales

Co-Op Details:

Able to work full time
Comfortable with AI concepts

Locations:

On-site at one of the Google offices:
Pittsburgh, Chicago, New York, Sunnyvale, San Diego, Las Angeles
We are very excited to see the potential impact you may have at Google.

Cheers,
Lucas Johnson
Google Recruiting



Note: We may use Google Meet to conduct and, in some cases, record interviews so that our team can focus on the conversation and review your responses accurately. Any recording will be kept confidential and used only for hiring-related evaluation purposes. You may request not to be recorded before the interview begins or ask to stop the recording at any time during the interview.

Google is committed to providing an inclusive and accessible interview experience for all candidates. If you need an accommodation at any stage of the interview process, please let your recruiter know, and we will work with you to provide appropriate support.

[Interviewer reads aloud] "Please take a look at this email. Assume you use Indeed."

A recruiter is interested in your profile Inbox x



Natalie <natalie@indeed-recruitment.com>
to me ▾

4:08 PM (42 minutes ago) ☆ 😊 ↶ ⋮

Hello! My name is Natalie from the Indeed Talent Strategy team.

We've identified your profile as a strong match for a specialized remote consultant role. This position involves assisting Tier-1 retailers such as Costco and Sam's Club with digital inventory optimization and brand exposure.

Role Details:

Time Commitment: Approximately 2-3 hours per day
Communication: \$25-30 per hour rate, settled daily via secure transfer.
Benefits: Full remote flexibility, 15 days of accrued PTO, and holiday pay.

We currently have 5 available positions for this quarter. If you would like to review the full job description and start the onboarding process, please reply to this email with your resume.



Section D: Review and Follow-up (2-3 min)

Here is where the interviewer can ask any follow up questions.

[Interviewer asks the following questions:]

- How confident are you in your ability to detect scams?
- Do you think that if you are stressed or anxious you might be more likely to act quickly without verifying information?
- Can you think of a time when your energy level affected how you responded to a situation?
- What specific designs or features do you look for when deciding whether or not an email is a scam?

I will now turn off the audio and video recording.

[turns off recording]

Section E: Debriefing and Wrap Up (1 min)

[Interviewer reads aloud:] “Thank you for participating in our study, ‘Understanding Interactions with Online Messages (STUDY2026_00000142).’ Our research focuses on how people understand and interact with online messages. Specifically we are looking at how people use motivated reasoning to interact with messages that could be scams. Some of the fictional messages we showed you were doctored versions of actual scam attempts. Just as a reminder, participation is totally voluntary and it is fully within your rights to ask for anything you said or any data to be withdrawn. Let us know now if you wish to withdraw anything you said, as we will be immediately de-identifying your responses.

Thank you for participating, do you have any questions or concerns? Please contact zsloneke@andrew.cmu.edu if you have questions at a later date. If you have questions or concerns about your rights as a research participant, you can contact the Carnegie Mellon IRB at (412) 268-4721, IRB-review@andrew.cmu.edu.

Once again, thank you. We will reach out within a week through email to send you your compensation.”