

Protecting Data Integrity: The Pervasive Challenge of Scammers in Remote Technology- based Research



Alexandra L. Terrill, PhD

Department of Occupational & Recreational Therapies

Department of Physical Medicine & Rehabilitation

Craig H. Neilsen Rehabilitation Hospital

DISCLOSURES

We have received funding from the National Institutes of Health NICHD/NCMRR under award number R03-HD091432 (PI: Terrill) and R01-HD105718 (PI: Terrill). We have also received funding from the Craig H. Neilsen Foundation (PSR Demonstration Project 720484, PI: Terrill).

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

OVERVIEW: OBJECTIVES

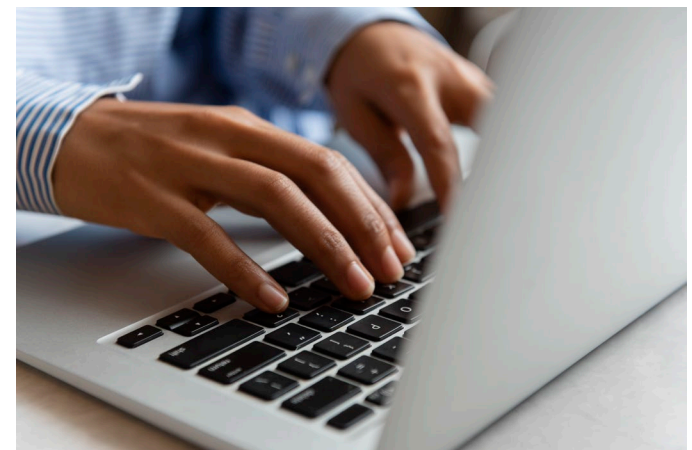
- Understand the threats posed by disingenuous responders when utilizing remote technology-based research design approaches.
- Learn about best-practice recommendations for identifying, addressing, and preventing scammers across different types of research designs.

OVERVIEW

- Real World Example Part 1: Bots and Insincere Responders in Surveys and Interviews (Cross-sectional Design)
- Real World Example Part 2: Randomized Control Trials (Longitudinal Design)
- Protecting Research Integrity: Best Practice Recommendations for Designing Studies
- Call to Action

REMOTE TECHNOLOGY-BASED RESEARCH

- What is it?
- Use of remote technology in research is growing
 - Rapid expansion in large part due to covid-19 pandemic



Nind et al., 2021

REMOTE TECHNOLOGY-BASED RESEARCH

Number of benefits

Reach

Accessibility

Time & Cost Effective

Participant Engagement

I think that it is **easier to share feelings** online. I think I would either not answer questions or I would give answers that were a little less honest if I were in a face-to-face group. In the online group, everyone is the same and you can't see anyone's reaction to what you say ([Nicholas et al., 2010](#), p. 116).

REMOTE TECHNOLOGY-BASED RESEARCH

But also Risks

Bots

Scammers

Insincere respondents



DO YOU USE REMOTE TECHNOLOGY-BASED APPROACHES IN YOUR RESEARCH?

Join at menti.com | use code 7909 2708

Mentimeter

Do you use remote technology-based approaches in your research?



13



Yes

0

No

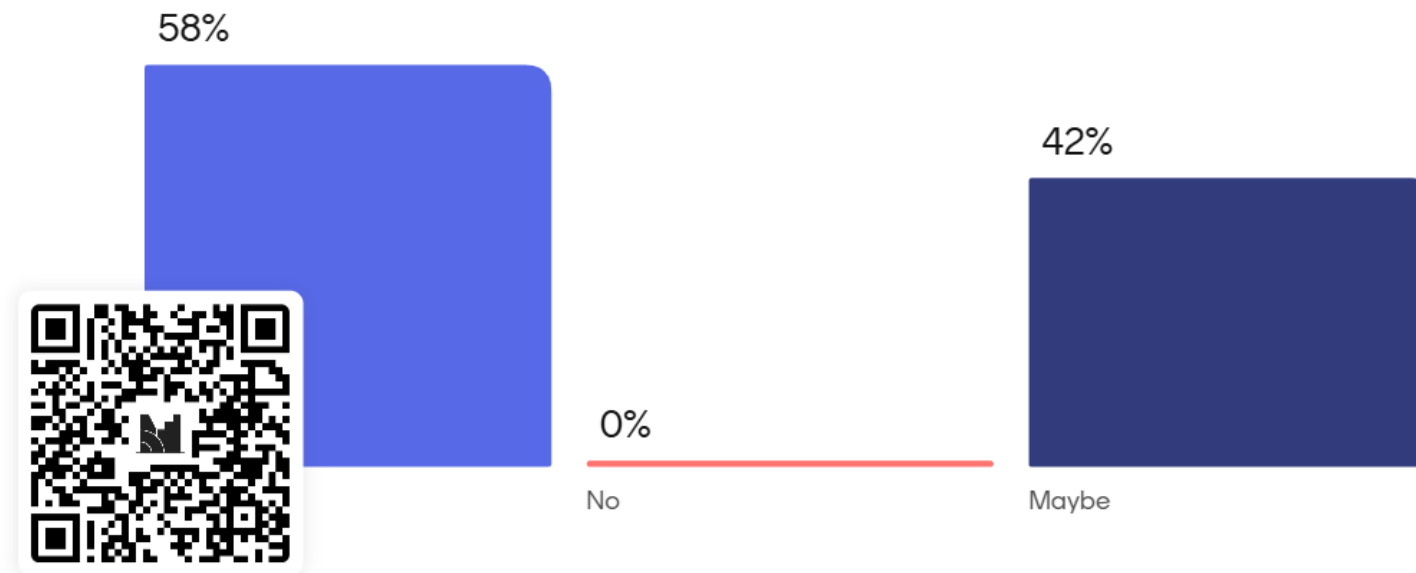
A vertical screenshot of a smartphone screen displaying the Menti poll interface. At the top, it shows the Menti logo and a user profile icon labeled 'AT'. Below this, the text 'Menti MAP' is visible along with share and refresh icons. A section titled 'Choose a slide to present' shows a list of four poll slides. Each slide contains the question 'Do you use remote technology-based approaches in your research?' and a progress bar. The first slide shows a blue bar for 'Yes' and a red bar for 'No'. At the bottom of the screen, there are icons for a thumbs up and a user count of 13.

HAVE YOU BEEN WORRIED ABOUT SCAMMERS IN YOUR RESEARCH?

Join at menti.com | use code 7909 2708

Mentimeter

Have you been worried about scammers in your research?



Real World Example Part 1:

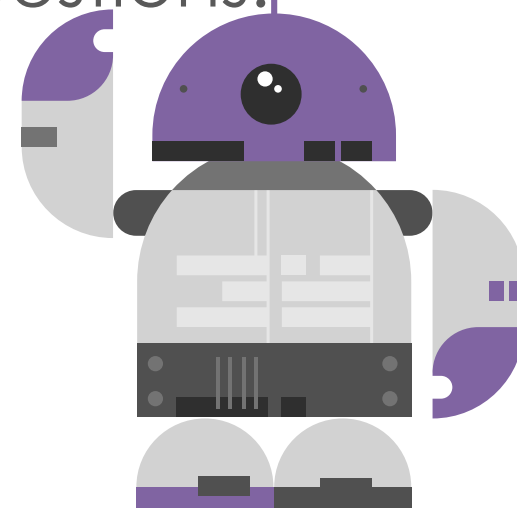
Bots and Insincere Responders in Surveys and Interviews (Cross-sectional Design)



Note: Portions of this presentation adapted from (and used with permission from)
Dr. Amber Thompson

INSINCERE & BOT RESPONSES IN ON-LINE SURVEYS

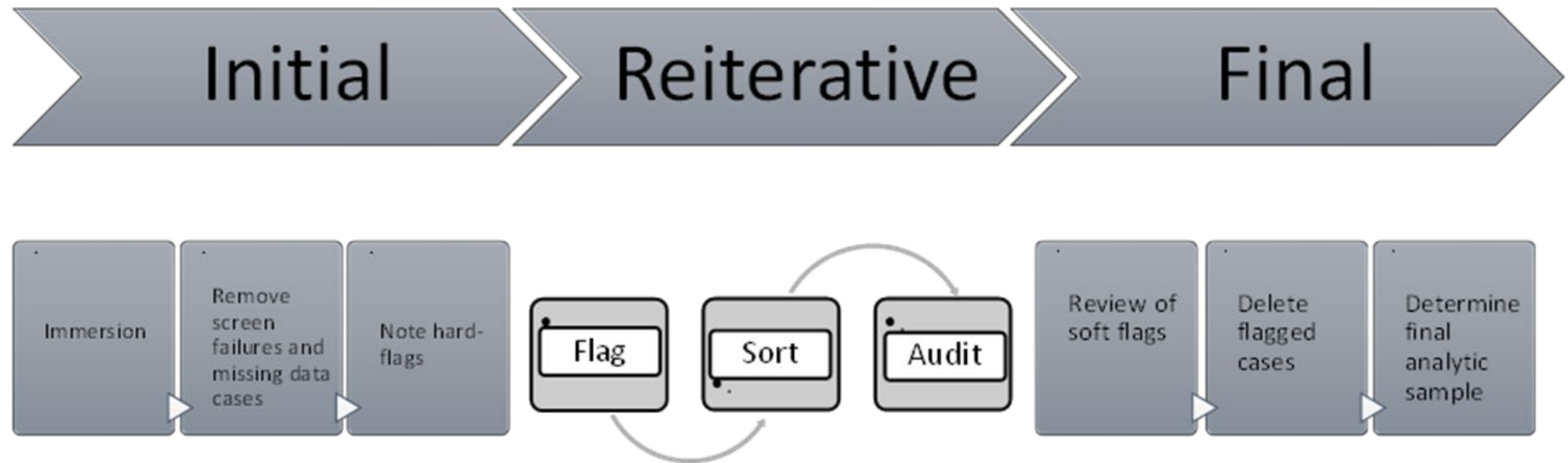
- **Bot response:** a simulated survey respondent that was likely submitted through an automated software or robot.
- **Insincere response:** responses that are not from a single individual (i.e., one person submits multiple responses) or is not from someone who is carefully and sincerely responding to survey questions.



INSINCERE & BOT RESPONSES IN ON-LINE SURVEYS: A CASE EXAMPLE OF A STUDY ON END-OF-LIFE DOULAS

- EOLD is a newly emerging role and a hard-to-reach population
- Online recruitment for web-based survey
- Protections were built in
- Both insincere and bot responses:
 - Produce data that is largely, if not completely, fabricated
 - Produce invalid data that compromises the sample and results
- Decisions had to be made based on available information and a systematic data-cleaning process was developed.

SYSTEMATIC REVIEW



Flags: hard versus soft flag based on whether the indicator could be used to *confirm* versus *suspect* possible data/cases to be dropped from the analysis.

Figure 1. Systematic Data Cleaning Process to Detect and Identify Insincere and Bot Responses on Web Surveys

PARADATA* CAN TELL US SOME THINGS

Start Date	End Date	Response Type	IP Address	Progress	Duration (in seconds)	Finished	Recorded Date	Location Latitude	Location Longitude
11/1/2022 21:26	11/1/2022 21:27	0	45.17.56.160	2	40	0	11/8/2022 20:27		
11/1/2022 21:26	11/1/2022 21:30	0	117.1.97.253	6	209	0	11/8/2022 20:30		
11/1/2022 21:29	11/1/2022 21:30	0	207.135.201.175	2	116	0	11/8/2022 20:30		
11/1/2022 21:28	11/1/2022 21:30	0	172.58.74.21	2	116	0	11/8/2022 20:31		
11/2/2022 2:02	11/2/2022 2:17	0	117.1.97.253	100	927	1	11/2/2022 2:17	21.0313	105.8516
11/3/2022 1:13	11/3/2022 1:37	0	117.1.97.253	100	1393	1	11/3/2022 1:37	21.0313	105.8516
11/3/2022 3:01	11/3/2022 3:39	0	117.1.97.253	100	2278	1	11/3/2022 3:39	21.0313	105.8516
11/30/2022 20:19	11/30/2022 20:47	0	117.1.111.186	100	1648	1	11/30/2022 20:47	21.0313	105.8516
12/1/2022 3:48	12/1/2022 4:03	0	117.1.111.186	100	870	1	12/1/2022 4:03	21.0313	105.8516
11/2/2022 2:01	11/2/2022 2:01	0	207.135.205.101	2	20	0	11/9/2022 1:01		
11/14/2022 2:21	11/14/2022 2:33	0	207.135.201.63	100	754	1	11/14/2022 2:33	28.2402	-82.7
11/1/2022 21:14	11/1/2022 21:22	0	172.58.251.84	100	484	1	11/1/2022 21:22	35.2296	-80.843
11/2/2022 17:34	11/2/2022 17:42	0	172.58.63.41	100	518	1	11/2/2022 17:42	39.7351	-105.027
11/1/2022 16:02	11/1/2022 16:03	0	172.58.204.248	6	90	0	11/8/2022 15:03		
12/4/2022 19:03	12/4/2022 19:06	0	172.58.73.25	27	174	0	12/11/2022 19:06		

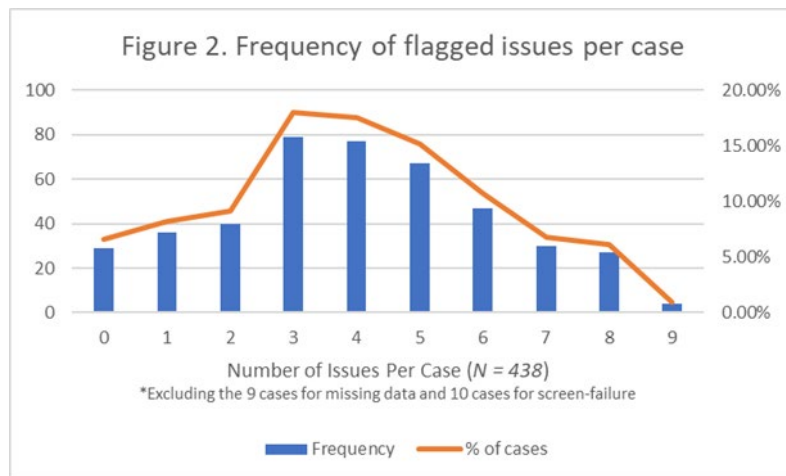
Start time, IP Address, % Complete, Duration in seconds, Location (latitude/longitude)

***paradata**: data about the process of data collection

QUALITATIVE DATA GIVES US SOME CLUES, TOO

Why did you become an EOLD? (Please describe below)	In your opinion, why is there a need for EOLDs? (Please describe below)
NEDA is committed to building a diverse and inclusive membership organization where communities can benefit from fair and equitable end-of-life care.	NEDA's mission is to influence how people experience positive change at the end of life through development and advocacy efforts to improve access to the wide range of holistic non-medical support offered by end-of-life doula.
æ°'ä¼—éœ€è 2	â°Šé+ç"Yâ'½
It's a good job to help others while getting rich pay.	People are helpless in the face of death, including himself and his family, so they need EOLD.
Like the service industry	Industry demand
Like the service industry	Social development needs
Can do more to help those ignorant people	Because many lonely old people are alone and helpless before they die, and they need help
Able to help others	Guide others
Want to be a guide to death	Everyone has no experience of death,so they need our guide
Want to be a witness of death	Death also requires ceremony
I like this industry	It can help them carry out orderly
Can serve others	Everyone is too busy with their work
Take care of others with a sense of achievement	Professional people do professional things
Help others	Social development needs
I need this job	social needs
Help people in need	Every family needs
Be happy for others to finish the last thing	Family needs
Like to take care of people	Everyone's needs
Help others to be supportive	Makes a person more satisfied when they end their life

FINDINGS FROM THE REVIEW



OUTCOME

94% of the respondents data were identified as having issues which prohibited their inclusion in the final analytic sample.

Table 2. Determining Final Analytic Sample Size, After Dropping Unusable and Invalid Cases by Tag

Case Drop-Reasons	Frequency in Full Sample (n=457)	Number of Dropped Cases for Specific Reason
Completely Missing	9	9
Screen Failures	10	10
Security Check & Auto-Generated Flags from Software Platform	2	2
Paradata Flags	295	276
Qualitative Responses	221	50
Quantitative Checks	234	25
Review of "Soft" flags	20	20
Incomplete Response; Partial Missing Data	104	36
Usable Cases	29	0

Note: A total of 457 surveys were submitted online via *Qualtrics*.

ILLUSTRATING THE IMPACT OF INSINCERE & BOT RESPONSES

Comparison of descriptive statistics using a parallel, directly recruited sample (n = 203)

*focus on characteristics reported in the literature

	Infiltrated Sample (n=402) <i>Mean or %</i>	Direct E-mail Sample (n = 203) <i>Mean or %</i>
How long have you been an EOLD?		
<12 months	23.65%	16.98 %
>1 <3 years	43.30 %	49.53 %
>3 > 5 years	23.36 %	18.40 %
>5 <10 years	7.69 %	10.85 %
>10 years	1.99 %	4.25 %
Is EOLD your primary occupation? --Yes	81.71 %	35.89 %
As an EOLD, do you get paid for support provided? --Always	51.30 %	14.08 %
Age (in years)	44	54
Gender—Female	51.09 %	91.46 %
Hispanic—yes	28.84 %	4.55 %
Race—Caucasian	52.80 %	87.06 %
Notes: Analytic sample sizes were determined by the total number of responses - screen failures – completely-missing responses – partially missing data that were missing data on key variables and would need to be practically dropped for analytic purposes.		

Table 3. Descriptive Statistics of Select Key Variables in Infiltrated and Direct-Email Samples

IT'S NOT JUST SURVEYS...

Qualitative Interviews with Families who Experienced EOLD Support:

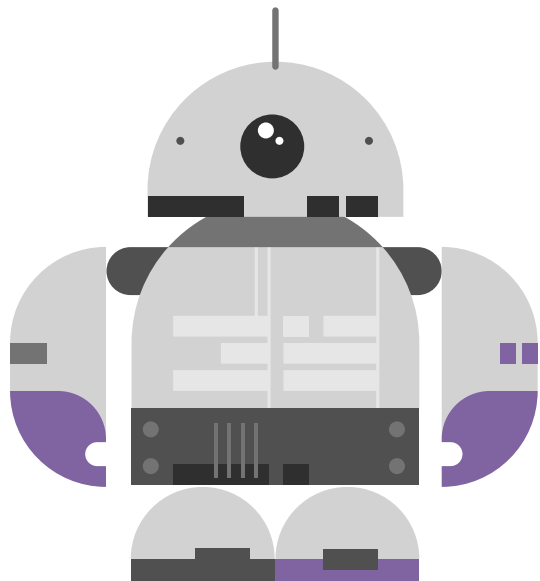
- Social Media Recruitment (N=7) and Zoom platform interviews.
- 6 identified insincere respondents

Issues:

- Patterns of shared characteristics: race, accent, age, background/location
- Other shared patterns: vague answers (or “textbook”), answers that contradicted known EOLD supports offered

A NEW WORLD IN RESEARCH

Bots can be sophisticated; insincere responses are a reality; and built-in platform protections may be inadequate in identifying either.



Lessons Learned:

Insincere and bot responses are:

- (1) becoming **bold** in their attempts to infiltrate research studies,
- (2) more **sophisticated** in their strategy to “pass” as sincere respondents,
- (3) unlikely to be easily detected by built-in platform security measures.

Data screening strategies: **Multiple strategies** for prevention and detection

Specific indicators of insincere or bot responses in surveys:

- Paradata
- Qualitative
- Quantitative

Specific indicators of insincere or bot responses in interviews:

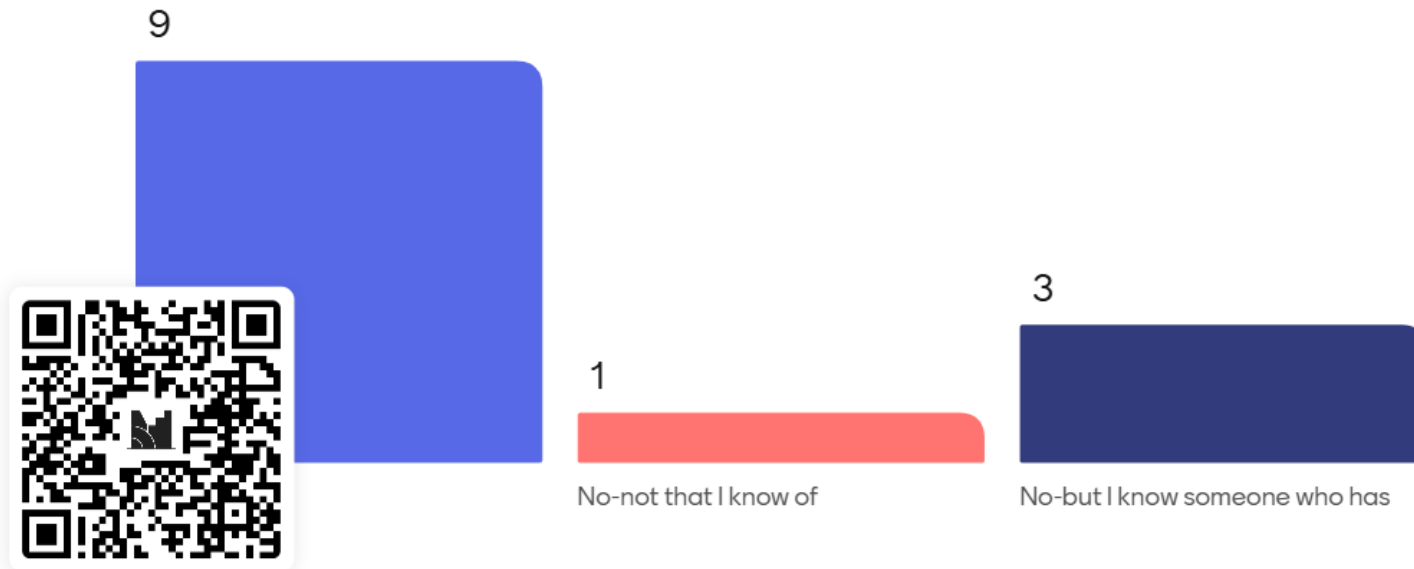
- Vague responses **and/or** little ability to expand
- Unexpected patterned characteristics of the target population

HAVE YOU EVER EXPERIENCED THIS ?

Join at menti.com | use code 7909 2708

Mentimeter

Have you experienced this kind of scamming/fraud?



FOR MORE INFORMATION:

Thompson, A.D., Utz, R.L. Online surveys: lessons learned in detecting and protecting against insincerity and bots. Qual Quant (2024). <https://doi.org/10.1007/s11135-024-01973-z>

Questions?

Real World Example Part 2: Randomized Control Trials (Longitudinal Design)

More sophisticated design = more sophisticated scammers





8-week remotely-delivered
intervention for couples
coping with stroke



8-week app-delivered
intervention for couples
coping with SCI



NIH R01-funded RCT (waitlist-control)

8 – week program:

- Reimagining Us in the Context of Stroke
- Goal Setting
- Self-care
- Communication
- Purpose and Meaning
- Connecting with Each Other
- Connecting with Others
- Looking to the Future

Assessments at weeks 1, 8, 16 & 40

All participation is remote



MODULE 4: LET'S TALK ABOUT IT



Communication allows us to share what we're thinking and feeling. Effective communication can also help other people understand what we need. It is easy to think that your partner may already know what you're thinking. Communicating well can help avoid misunderstandings, resolve conflict, and make our partners, family, and friends feel appreciated, deepening those relationships.

It can be challenging to start some conversations, especially about personal or sensitive topics. A stroke can also affect communication, making it even more difficult to share thoughts with others. However, do not get discouraged if you find it awkward or frustrating. Communication is hard work, but gets easier with practice.



ACT OF KINDNESS is a selfless act performed by kind people to either help or cheer up other individuals, for no reason other than to make another person happier. It is giving for the sake of giving, not in order to receive something in return.

EXAMPLE ACTIVITIES:

Why do **ACTS OF KINDNESS** make you happier, healthier, and live longer? An act of kindness relieves stress and improves mood by distracting from one's own troubles, has a positive effect on how we view ourselves, enhances sense of meaning and purpose in life, enhances social integration, and strengthens social bonds.

Try these activities by yourself:

1. Share kind thoughts with a friend or neighbor. Send a text, email, letter, or call and tell the person about their strengths and positive characteristics.
2. Donate unwanted clothes/supplies to homeless or rescue shelters.
3. Donate to a local charity if you are able or share information about a charity on social media to help promote and spread the word about the organization.
4. Practice self-kindness by asking for help with tasks when you need it. As a caregiver, this can mean asking for relief from caregiver duties for even 15-30 minutes.



Support Groove

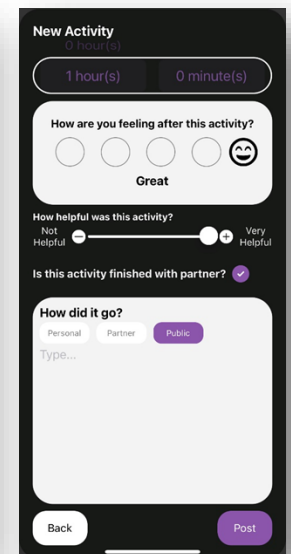
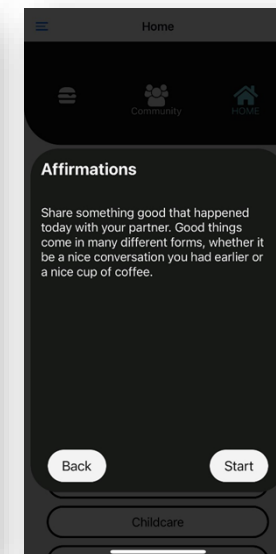
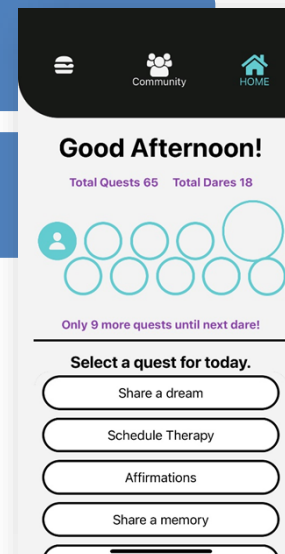
Foundation-funded RCT (waitlist control)

8 – week program:

- Uses an app
- Each week, receive a themed activity prompt

Assessments at weeks 1, 8, 16 & 28

All participation is remote





Inclusion Criteria

- Must be at least 18 years old
- 3 months to 3 years out from most recent stroke
- Must have a romantic partner (living together)
 - both partners must be willing to participate
- Ability to read and follow instructions in English

Exclusion Criteria

Significant cognitive impairment / aphasia (unable to provide own consent)

Notes:

Mild to moderate cognitive impairment and expressive aphasia okay

Enrollment in other studies okay

This is **NOT** couples therapy



Inclusion Criteria

- Must be at least 18 years old
- At least 3 months out from injury
- Must have a romantic partner (living together)
 - both partners must be willing to participate
- Ability to read and follow instructions in English
- Have internet access & a tablet or smart phone

Exclusion Criteria

A major neurological condition in the partner

Something's amiss...



THE “FAKE COUPLE”

- First dyads of the study
- Single person playing both roles
- Enrolled as multiple dyads
- Same Apple ID
- Same IP address



TAKING ACTION

 Suspend enrollment for BOTH studies

 Contact IRB

 Contact funding agency

Thank you for letting me know about this issue.

We are hearing about similar issues, such as bots and other Internet-specific dynamics, from grantees. Sadly, I am not surprised.

TAKING ACTION



Develop and implement “**anti-scammer protocol**” for BOTH studies

- Changes to study procedure
 - “I understand I may be required to provide proof of identity in the form of a US government-issued photo ID prior to enrollment.”
 - Visually inspecting photo ID via video call
- Change to consent form
 - The study investigators...
 - Best interest
 - Do not follow rules/procedures
 - Provide false information
 - Experience serious side effects

ANTI-SCAMMER IN REAL TIME

- What do we look for when we first see referrals/self-referrals?
 - Red flags
- What changed in the way we screen?
 - Ongoing caution / scrutiny
- Our informed consent revision
 - Allows us to drop if scammer enrollment happens (“Force quits”)

“I am interested in your research.”

“Hello, With great delight I am pleased to be enlightened more about this study and will like to be reached to take part in the study.”

“I recently came across information about the study you are conducting and am very interested in participating. I believe I meet the specified criteria:

1. I experienced a stroke at least 3 months ago;
2. I live with my romantic partner who is also willing to participate with me;
3. We are both over 18 years old.”

ANTI-SCAMMER IN REAL TIME

- Referral sources
 - ResearchMatch
 - University Study Locator

“John Doe” / 24M

Email: doejohn030@gmail.com (name + numbers is common scammer email)

Conditions: **alzheimer disease**, brain tumor, **type 2 diabetes**, food allergies, **heart disease**, high blood pressure, leukemia, **stroke**, uti - urinary tract infection, chronic pain, sexual abuse, thyroid cancer, severe depression, lung cancer

EXAMPLES OF SUSPICIOUS REFERRALS

- “Jason Greg” / 41M
- Conditions: aids, anemia, anxiety, anxiety disorder, joint pain, arthritis, psoriatic arthritis, rheumatoid arthritis, asthma, exercise induced asthma, astigmatism, atrial fibrillation, bipolar disorder, bipolar depression, colon cancer, skin cancer, colitis, ulcerative colitis, diabetes type 1, type 2 diabetes, eczema, **endometriosis**, epilepsy, fatigue, chronic fatigue syndrome, fatty liver, fibromyalgia, congestive heart failure, hiv, hiv positive, high blood cholesterol level, hypertension, hypothyroidism, kidney disease, copd (chronic obstructive pulmonary disease), nausea, degenerative joint disease, parkinson's disease, psoriasis, kidney failure, sinusitis, stomach ulcer, stress, ptsd - post-traumatic stress disorder, **stroke**, syphilis, chronic sinusitis, plantar fasciitis, migraines, chronic pain, allergic asthma, covid-19, **post partum depression**, knee joint pain, bipolar ii disorder, generalized anxiety disorder, depression with anxiety, prediabetes, **prostate cancer**, chronic kidney disease, knee osteoarthritis, **enlarged prostate**, neuropathy, kidney stone passed, functional state, thyroid cancer, sleep problem, chronic back pain, chronic insomnia, kidney problem, bipolar i disorder, insomnia, adhd (attention deficit hyperactivity disorder), major depressive disorder, hypermobile joints

- “Max Harold” / Email: haroldmax67@gmail.com

ReStoreD: Sent standard introductory email *(phone number was not provided)*

Max: I am comfortable with the criterias. We can make a Teams call tomorrow from 2P.M *(seems fishy, but he is offering a video call – so we continued)*

ReStoreD: I'm glad to see you are interested in learning more about the ReStoreD study. Unfortunately, I am not available tomorrow at 2:00. I have 9:00am or 4:00pm available. These times are Mountain Standard Time. Please let me know if either of these times will work for you and your partner.

Will you please also provide your partner's name and a good phone number to reach you?

Max: I am available by 9A.M tomorrow for a zoom or teams call *(still no partner info)*

ReStoreD: Thank you. I do need your partner's name and a phone number to schedule the call. Please confirm that your partner will be joining the call and provide this additional information.

Max: My partner is comfortable joining the call *(still no partner info)*

ReStoreD: Please call the number listed below to coordinate a time for me to review the study with you and your partner. I will need to gather some additional information before I can schedule a zoom call.

End of communication – moved to not enrolled / scammer

Was this a scammer?

ANTI-SCAMMER IN REAL-TIME

- Use of AI during screening calls



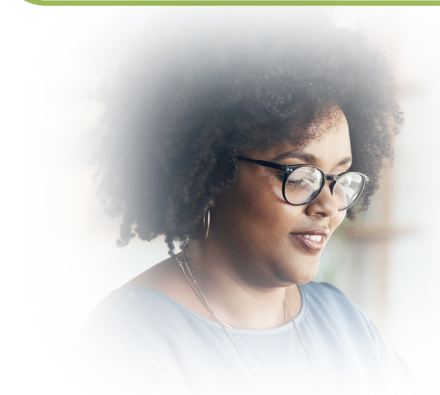
How do I answer the following question as if I were a stroke survivor: "How has your stroke affected you?"

"What kinds of changes in your mood have you noticed since your stroke?"

Here's a possible way to answer:

"My stroke has affected me in profound ways, both physically and emotionally. On a physical level, I faced challenges like mobility issues, speech difficulties, and fatigue. It's been a journey of adapting to these changes and finding new ways to do things that once came effortlessly."

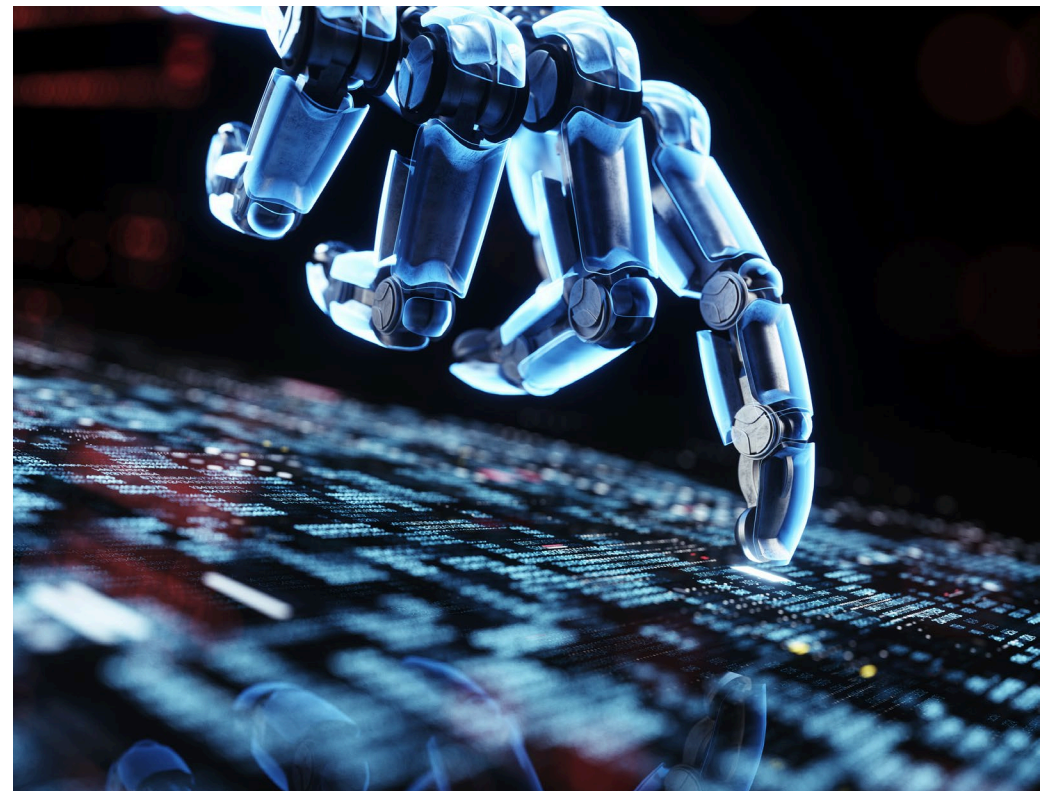
"My mood has changed a lot since the stroke. Sometimes I feel sad, frustrated, or even angry because of the challenges I face. It can be hard to accept the changes and keep going."



ANTI-SCAMMER IN REAL TIME

- How often is this happening?
 - About 10% - 12% of referrals
 - more than half of referrals we get from ResearchMatch and most of the Study locator referrals are easily identified as scammers
 - Repeat offenders

QUESTIONS?



Protecting Research Integrity: Best Practice Recommendations for Designing Studies



ETHICAL CONSIDERATIONS: INCREASING PREVALENCE AND SOPHISTICATION OF FRAUD

- Rates:
 - 18% - 35% fraudulent responses on surveys
 - 24% - 83% in surveys posted to mTURK on Amazon (Wessling et al., 2017)
 - 23.5% in a fully automated 8-wk Internet-based CBT RCT (Schure et al., 2019)
 - >7% in a 62-wk clinical trial with multiple assessments, including phone interview (Glazer et al., 2021)

Impact on outcomes & conclusions drawn from research are significant

Ballard et al., 2019; Schre et al., 2019; Glazer et al., 2021; Wessling et al., 2017; Young et al., 2020

WHO ARE THESE SCAMMERS?

- Ranges from “one-offs” to large-scale warehouse operations



'Hundreds of thousands' trafficked into SE Asia scam centres - UN

By Reuters

August 29, 2023 4:14 AM MDT

Cyber Scamming Goes Global: Unveiling Southeast Asia's High-Tech Fraud Factories

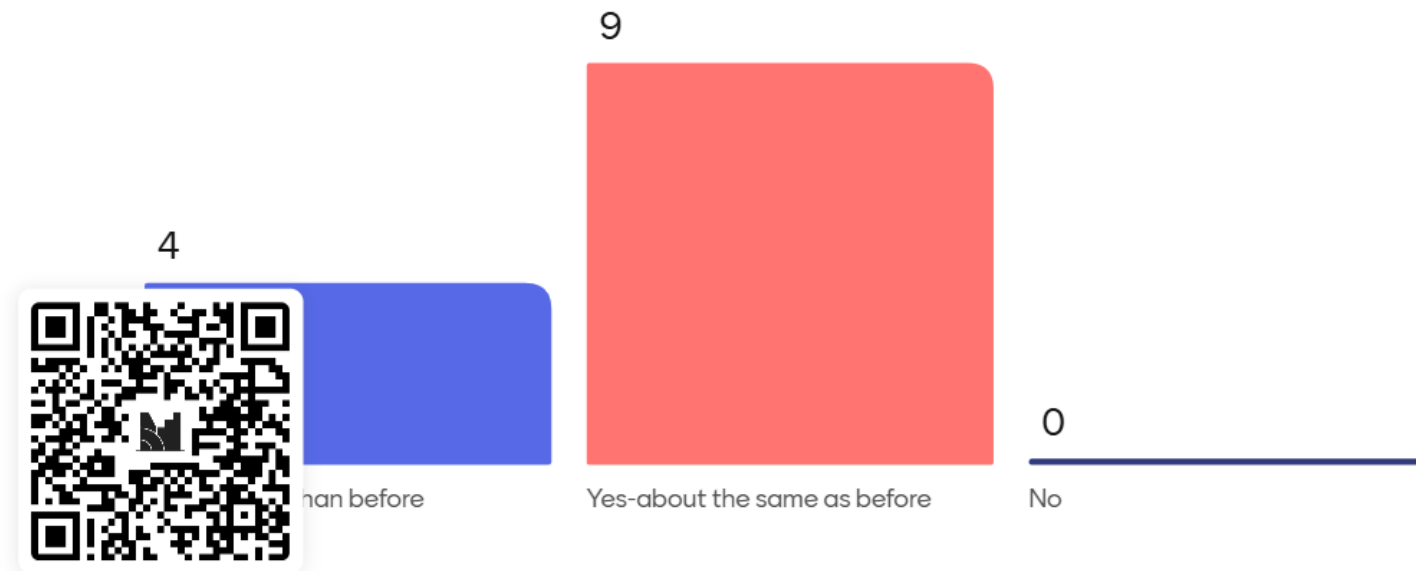
Wikipedia: A scam center, fraud factory, fraud park, scam factory, scam compound, scam hub, scam park, fraud center, fraud compound, or fraud hub is a collection of large fraud organizations usually involved in human trafficking operations, generally found in Southeast Asia and usually operated by a criminal gang.

ARE YOU WORRIED ABOUT SCAMMERS IN YOUR RESEARCH?

Join at menti.com | use code 7909 2708

Mentimeter

Are you worried about scammers in your research?



What can be done?



THE BALANCE

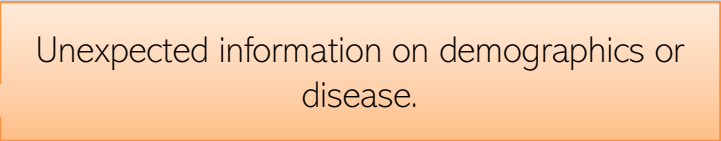
- Protecting participants' confidentiality to encourage them to trust the researchers
- Making research more accessible
- Ability to reach more diverse, harder-to-reach participants
- Providing incentives / compensation for participant time
- Researchers' ability to trust the participants
- Ensuring data integrity
- Incentives increases risk 6x



BEST PRACTICE RECOMMENDATIONS BASED ON EVIDENCE AND EXPERIENCE

The ethical issues and IRB considerations involved have received little systematic attention

DETECTING & PREVENTING FRAUD: INSTRUMENT

Type of intervention	Method of Prevention	Pros	Cons	Additional ethical concerns
Questions : Inconsistent responses	Check for consistency	Can detect 	Skipped questions	
Include similar/strange question	Include similar or strange question	Level of attention	Can impact experimental design	
Software	No back button	Can't easily resubmit survey	No prevention Limits participants from going back	
	Change order	Can detect bots		
	CAPTCHA	Detects bots		
	Collect paradata	Response patterns	Costly	Should this be disclosed to participants?

Adapted from Glazer et al., 2021; Mistry et al, 2024; & Teitcher et al., 2015

DETECTING & PREVENTING FRAUD: NON-INSTRUMENT DATA

Type of intervention	Method of Prevention	Pros	Cons	Additional ethical concerns
Personal information: Similar/same email b/w participants	Contact partic., if no response, drop	Clears up misunderstanding	Response bias Deter eligible partic. Doesn't stop dissimilar	Need to balance protecting integrity of data and subject privacy and confidentiality are particularly important
Inaccurate/ fake address, phone or email	Request phone #/adress	Need valid # to proceed	Can create temporary phone #s	
Verify identity	Check if pe address, phone is valid	Consistent info Deters fraud	Can discourage eligible participants	
	Check info against external sources Identity verification service Valid ID	Confirms for consistent info Deters fraud	Can discourage eligible participants Verification service can be pricey Technological barrier for participants	

Email address format:
johndoe999@gmail.com
doejane678@gmail.com
jimjohn452@gmail.com

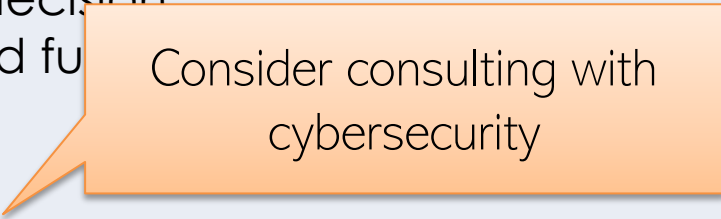
DETECTING & PREVENTING FRAUD: COMPUTER INFORMATION

Type of intervention	Method of Prevention	Pros	Cons	Additional ethical concerns
IP address: Same IP address b/w participants	Check if IP address is same or encrypted	Can determine how many times survey taken Location criteria	IP address may be shared IP address can be scrambled/ encrypted Advanced programs to check rerouting are costly	Privacy issue (is IP address personal info?) Should consent forms mention IP address tracking?
	Block IP address if fraud	Avoids fraud	Could be dynamic (but not fake)	
Internet cookies: cookies detect browser, device information	Enable cookies (e.g., Qualtrix "ballot-box stuffing")	Can detect multiple submissions	Can disable cookies Can use different browsers Computers may be shared	Administer cookies without participant knowledge Should they be informed?

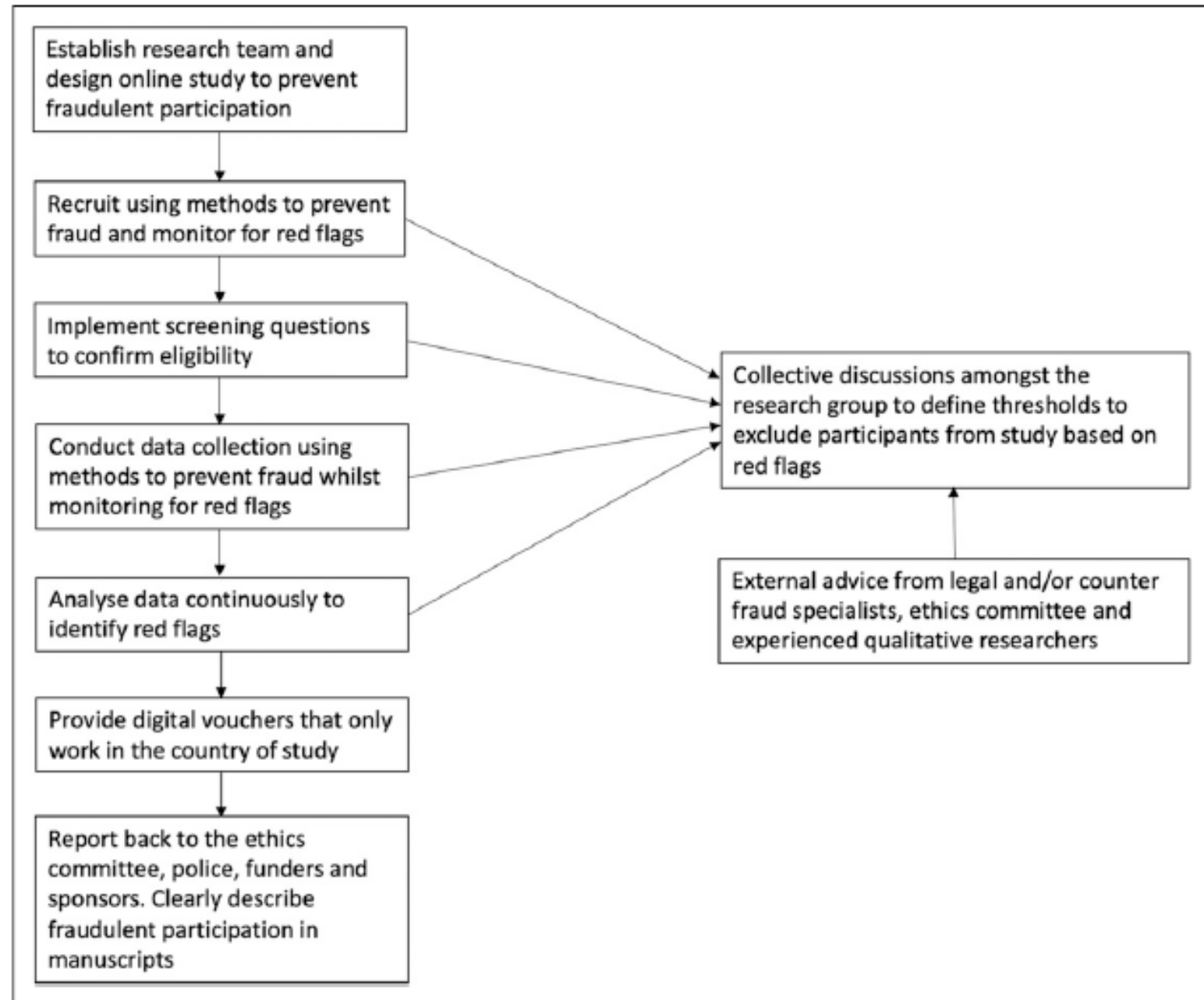
DETECTING & PREVENTING FRAUD: STUDY DESIGN

Type of intervention	Method of Prevention	Pros	Cons	Additional ethical concerns
Informed consent	Break up consent online, online provide payment info at end	Attention needed, longer process. Discourages fraud Reduces bots	May deter eligible participants	Needs to balance protecting integrity of data and subject privacy and confidentiality are particularly important
Secondary screen	Phone screen / consent review over video/phone (ID check)	Deters fraud	May deter eligible (diverse) participants	
Compensation	Mention no compensation if suspected of fraud	Avoids paying fraud Consider not advertising compensation		
	Pay contingent on verifying ID & mailing address	Deters ineligible participants	May deter eligible participants	
	Gift card lottery	Gives researchers time to review before compensation	May not be enough to incentivize eligible participants Fraud take more surveys	
Including interview	Audio or video interview	May deter fraud Can detect lying	Fraud can be hard to detect Lose anonymity; may deter participants	

DETECTING & PREVENTING FRAUD: BROADER REGULATORY BOARDS

Type of intervention	Method of Prevention	Pros	Cons	Additional ethical concerns
IRBs: IRB structure	Have online/computer expert as member	Ability to assess study at hand and find balance to protect participants and ensure data quality	Does not deter fraud	
PIs report fraud to IRB	IRB can make appropriate decision for current and future studies	May deter fraud 	May deter legitimate participants Can create new names, emails, etc	
Broader entities: report information	PIs create fraud list and share with others Report to Internet Crime Complain Center (IC3.gov) or funder	List can help identify and deter fraud May deter fraud	Can create new names, emails, etc May deter legitimate participants	

EXAMPLE STANDARDIZED APPROACH (P-FROST)



Mistry et al., 2024

KEY POINTS:

- Establish a protocol up front that includes:
 - Thorough screening and verification
 - Check the (para)data
 - Establish “Force Quits”
 - All team members should be trained on protocol
- Trust your gut
- Follow the flags



A Call to Action



A CALL TO ACTION: PRINCIPAL INVESTIGATORS

- Primary responsibility is on the PI
- BUT there may be a reluctance to discuss
 - Question findings (current and past)
 - Ability to publish
 - Worry about loss of/future funding

A CALL TO ACTION: PRINCIPAL INVESTIGATORS

Online Survey Fraud Detection Algorithm

Each entry was initially reviewed using the following criteria and assigned a fraud score (points given in brackets):

- a) [1 point] Was the IP address location outside of the study area AND unable to be verified through other means (ie, a mailing address for mailed incentives AND completion in-person at a study recruitment event)? (n=164)
- b) [1 point] Is the provided phone number a business or organization number rather than personal number? (n=94)
- c) [1 point] Are names within the entry (ie, in consent form, email address, incentive mailing address) different? (n=43)
- d) [1 point] Is the email address provided unusual (ex. a12bcd34e@email.com)? (n=37)
- e) [2 points] Does the first and last name AND one or more other personal information items (ie, phone number, email address, and/or physical address) match previous entries'? (n=13)
- f) [1 point] Do two or more of the following match previous entries' information: phone number, email address, and/or physical address? (n=6)
- g) [2 points] Does the date of birth AND one or more personal information items (ie, phone number, email address, and/or physical address match previous entries'? (n=5)
- h) [1 point] Was the duration of the survey less than 5 minutes? (n=3)

All entries that received a score of 1 point were categorized as potentially fraudulent. All entries that received a score of 2 points or more were categorized as fraudulent.

Table 1. Indicators of Scams.

Indicators During the Recruitment Process

- ☐ Large volume of emails received in a short period of time
- ☐ Timing of Email receipt
- ☐ Similar email address format
- ☐ Similarity in the message contents
- ☐ Rapidity of response

Indicators During the interviews

- ☐ Refusal to turn the camera on
- ☐ Poor sound quality, technical problems and other sound cues
- ☐ Interview duration
- ☐ Brevity of responses
- ☐ Incoherence and inconsistencies
- ☐ Scammers refer to similar content
- ☐ Strong financial interests

A CALL TO ACTION

- Primary responsibility is on the PI
- Funding agencies
- Journal editors
- Reviewers

**Rigor &
Transparency**



An aerial photograph of the University of Utah Health building, a large multi-story structure with a modern architectural style featuring glass and brick. The building is surrounded by a landscaped courtyard with trees, walkways, and a few parked cars. A red vertical bar is visible on the far left edge of the image.

THANK YOU



HEALTH
UNIVERSITY OF UTAH

REFERENCES

- Ballard, A. M., Cardwell, T., & Young, A. M. (2019). Fraud Detection Protocol for Web-Based Research Among Men Who Have Sex With Men: Development and Descriptive Evaluation. *JMIR Public Health Surveill*, 5(1), e12344. doi:10.2196/12344
- Bowen, A. M., Daniel, C. M., Williams, M. L., & Baird, G. L. (2008). Identifying Multiple Submissions in Internet Research: Preserving Data Integrity. *AIDS and Behavior*, 12(6), 964-973. doi:10.1007/s10461-007-9352-2.
- Glazer, J. V., MacDonnell, K., Frederick, C., Ingersoll, K., & Ritterband, L. M. (2021). Liar! Liar! Identifying eligibility fraud by applicants in digital health research. *Internet Interventions*, 25, 100401. doi:https://doi.org/10.1016/j.invent.2021.100401
- Mistry K, Merrick S, Cabecinha M, Daniels S, Ragan J, Epstein M, Lever L, Venables ZC, Levell NJ. Fraudulent Participation in Online Qualitative Studies: Practical Recommendations on an Emerging Phenomenon. *Qual Health Res*. 2024 Nov 16:10497323241288181. doi: 10.1177/10497323241288181.
- Nind, Coverdale, & Meckin (2021) Changing Social Research Practices in the Context of Covid-19: Rapid Evidence Review. Project Report. NCRM.
- Pullen Sansfaçon, A., Gravel, E., & Gelly, M. A. (2024). Dealing With Scam in Online Qualitative Research: Strategies and Ethical Considerations. *International Journal of Qualitative Methods*, 23, 16094069231224610. doi:10.1177/16094069231224610
- Roberts, J., Onuegbu, C., Harris, B., Clark, C., Griffiths, F., Seers, K., Aktas, P., Staniszezwska, S., & Boardman, F. (2025). Comparing In-Person and Remote Qualitative Data Collection Methods for Data Quality and Inclusion: A Scoping Review. *International Journal of Qualitative Methods*, 24. https://doi.org/10.1177/16094069251316745
- Schure, M. B., Lindow, J. C., Greist, J. H., Nakonezny, P. A., Bailey, S. J., Bryan, W. L., & Byerly, M. J. (2019). Use of a Fully Automated Internet-Based Cognitive Behavior Therapy Intervention in a Community Population of Adults With Depression Symptoms: Randomized Controlled Trial. *J Med Internet Res*, 21(11), e14754. doi:10.2196/14754.
- Teitcher, J. E., Bockting, W. O., Bauermeister, J. A., Hoefer, C. J., Miner, M. H., & Klitzman, R. L. (2015). Detecting, preventing, and responding to "fraudsters" in internet research: ethics and tradeoffs. *J Law Med Ethics*, 43(1), 116-133. doi:10.1111/jlme.12200
- Thompson, A.D., Utz, R.L. Online surveys: lessons learned in detecting and protecting against insincerity and bots. *Qual Quant* (2024). https://doi.org/10.1007/s11135-024-01973-z
- Wessling, K. S., Huber, J., & Netzer, O. (2017). MTurk character misrepresentation: Assessment and solutions. *Journal of Consumer Research*, 44(1), 211-230. doi:10.1093/jcr/ucx053.