Tip Sheet: Managing Bots and Imposter Participants in Research

As researchers at the University of Alberta, you are dedicated to generating high-quality, reliable research and evidence to inform health practices and policies. Maintaining the integrity of your research is paramount. This tip sheet addresses a growing challenge: the presence of bots and imposter participants, and offers practical strategies to mitigate their impact on studies.

The Problem

Recruiting participants is a critical step in ensuring the success and validity of your research study. The increasing accessibility of online platforms for research recruitment offers numerous benefits but also introduces the risk of encountering non-genuine participants, including bots and imposter participants.

- **Bots:** Automated software programs that mimic human responses. In research, they can be programmed to complete surveys, click on links, or even attempt to engage in online data collection.
- **Imposter Participants:** Individuals who intentionally misrepresent themselves or provide false information to participate or gain incentives without meeting the eligibility criteria.

Both issues can compromise data quality and the integrity of research findings and result in wasted resources. In addition, maintaining trust in research findings is crucial for their translation into practice. As such, it is essential to implement strategies to identify and mitigate the risks posed by bots and imposter participants throughout the research process.

Strategies for Identifying Bots and Imposter Participants

Here are actionable strategies to minimize the risk of bot and imposter participant involvement in your research. **The appropriate strategies you use will vary** depending on your study methodology (i.e., qualitative, quantitative, or mixed/multi-methods), study population (e.g. vulnerabilized populations), and sampling strategy.

1. Identifying Bots

- **CAPTCHA and reCAPTCHA:** Implement these challenges at the beginning or critical junctures to differentiate between humans and automated programs.
- **Time-Based Measures:** Consider unnaturally fast completion times to indicate bot activity. However, be mindful of accessibility needs and the potential for genuine fast responders.
- **Pattern Recognition Questions:** Embed questions designed to detect consistent, non-human responses (e.g., a honeypot question or a series of Likert scale questions where the same option is selected repeatedly or a question with an obvious answer).
- **Generic answers:** Analyze open-ended responses for nonsensical, irrelevant, or copied text. Bots are increasingly capable of generating text, so this is not a foolproof method.
- **IP Address and Geolocation Analysis:** While respecting privacy, monitoring for multiple submissions from the same IP address or unusual geographic locations can flag suspicious activity. Be cautious as multiple participants might share an IP address (e.g., within an institution). Ensure you have appropriate research ethics approval for collecting and/or using this type of identifiable data.

2. Identifying Imposter Participants

- **Pre-Screening Questionnaires:** These questionnaires collect information that is not used in a data set. Ask closed- and open-ended questions to assess eligibility and genuine interest. Look for inconsistencies or red flags in their responses.
- **Confirmation Calls/Emails:** Contact potential participants to verify their identity and interest before scheduling the session. When possible, contact participants by telephone as this helps identify general locations (through area codes) and the participant's voice (to identify duplicate imposter participants).
- **Video Conferencing:** When conducting online interactions, requiring participants to turn on their video ***during screening and before data collection*** can help verify their identity and increase accountability. If appropriate to your study methodology and population, you might consider having potential participants show valid identification with their name, photo, and/or address during screening, depending on your eligibility criteria.
- **Attention Check Questions:** Integrate simple questions to assess participant engagement (e.g., "To show you are reading carefully, please select 'Strongly Disagree' for this question" or "write 44 in the chatbox").
- **Time-Based Measures:** Consider unnaturally fast completion times to indicate imposter participant activity. Again, be mindful of accessibility needs and the potential for genuine fast responders.
- **Inconsistency Checks:** Include related questions phrased differently to identify contradictory responses or include duplicate questions to check for consistency. Use your judgment, as eligible participants may genuinely contradict themselves based on their circumstances, or such contradictions could be interesting findings in qualitative research. Analyze open-ended responses for nonsensical, irrelevant, or copied text.
- **Generic or vague answers:** Analyze open-ended responses for nonsensical, irrelevant, or copied text.
- **Lack of Engagement:** Participants who are passive, unresponsive, or seem uninterested in the interaction could be non-genuine. Participants providing overly general or non-specific answers without personal details or insights might be suspicious. Be attentive to responses that contradict earlier statements or do not make logical sense.

- **Suspicious Background Noise or Disruptions:** While not always indicative of fraud, unusual or disruptive background noises during online sessions might warrant further attention.

Remember to use your judgment and consider the needs of your study methodology and study population when deciding which strategies to use and when to use them.

Strategies for Managing Bots and Imposter Participants

Here are actionable strategies to mitigate the impact of bots and imposter participants. Remember, the appropriate strategies will vary depending on your study methodology (i.e., qualitative, quantitative, or mixed/multi-methods), study population (e.g. vulnerabilized populations), and sampling strategy.

- **Ethics Considerations:** List all of the strategies you will use to identify and manage bots or imposter participants in your research ethics proposal. Ensure all screening and verification methods comply with ethical guidelines, respect participant privacy, and are approved by the Research Ethics Board at the University of Alberta.
- **Eligibility Criteria:** Clearly define and adhere to your study's inclusion and exclusion criteria.
- **Multi-Stage Screening:** Implement a layered approach to screening participants at different stages of the research process (e.g., initial application, pre-participation questionnaire, during data collection). Consider phone or video screening for qualitative research.
- **Excluding a potential participant due to suspected fraud:** Use clear, direct language: (e.g. "Thank you for your time. We've determined that you do not meet the study criteria, and therefore, we are unable to include you. Goodbye.")
- **Survey Design Elements:**
    - **Logical Flow and Skip Logic:** Ensure tasks flow logically and use skip logic to avoid irrelevant questions, which can confuse genuine participants and potentially flag them incorrectly.
    - **Clear and Concise Language:** Use straightforward language to minimize misinterpretation.
    - **Progress Indicators:** Help maintain participant engagement and reduce the likelihood of rushed or random responses.
- **Pilot Testing:** Conduct a pilot study with a small group to identify potential issues with your recruitment methods and data collection instruments, including the strategies to reduce bots or suspicious participants. Consider whether this data will be included in your data analysis.
- **Data Monitoring:** Continuously monitor incoming data for anomalies, inconsistencies, and patterns that might indicate fraudulent activity.
- **Data Validation Techniques:**
    - Statistical Outlier Analysis: Identify responses that fall outside the expected range for certain variables.
    - Mahalanobis Distance: A multivariate outlier detection technique that considers the correlation between variables.

- **Intuition and Professional Judgment:** Trust your instincts. If a participant's behaviour or responses seem "off," document your concerns and consider excluding their data. Discuss these instances with your research team.
- **Documentation:** Maintain detailed records of your screening processes and any instances of suspected fraudulence, including the rationale for exclusion of the participant and/or the data collected.

Quick Look Table

In the **Quick Look Table**, we provide a list of strategies to identify and manage bots and imposter participants at different research stages.

RESEARCH STAGE	BOTS	IMPOSTER PARTICIPANTS
Planning (ethics application, survey/question development)	• CAPTCHA and reCAPTCHA • Eligibility Criteria • Ethics Considerations • IP Address and Geolocation Analysis • Multi-Stage Screening • Pattern recognition questions • Pilot Testing • Survey Design Elements	• Confirmation Calls/Emails • Eligibility Criteria • Ethics Considerations • Inconsistency Checks • IP Address and Geolocation Analysis • Multi-Stage Screening • Pilot Testing • Pre-Screening Questionnaires • Survey Design Elements • Video Conferencing
Recruitment and Screening	• CAPTCHA and reCAPTCHA • Eligibility Criteria • Multi-Stage Screening • Excluding a potential participant due to suspected fraud • IP Address and Geolocation Analysis	• Confirmation Calls/Emails • Eligibility Criteria • Excluding a potential participant due to suspected fraud • IP Address and Geolocation Analysis • Multi-Stage Screening • Pre-Screening Questionnaires • Video Conferencing

RESEARCH STAGE	BOTS	IMPOSTER PARTICIPANTS
Data collection	• Data Monitoring • Excluding a potential participant due to suspected fraud • Survey Design Elements • CAPTCHA and reCAPTCHA • IP Address and Geolocation Analysis	• Attention Check Questions • Data Monitoring • Excluding a potential participant due to suspected fraud • Inconsistency Checks • Lack of Engagement • Suspicious Background Noise or Disruptions • Video Conferencing
Data cleaning	• Data Monitoring • Data Validation Techniques • Generic answers • IP Address and Geolocation Analysis • Time-Based Measures	• Data Monitoring • Data Validation Techniques • Generic or vague answers • IP Address and Geolocation Analysis • Suspicious Background Noise or Disruptions • Time-Based Measures
Data analysis	• Data Validation Techniques • Generic answers • IP Address and Geolocation Analysis	• Data Validation Techniques • Generic or vague answers • Inconsistency Checks • IP Address and Geolocation Analysis • Lack of Engagement • Suspicious Background Noise or Disruptions

Remember to use your **intuition and professional judgment** at every stage of research. Consider the needs of your study methodology and study population when deciding which strategies to use and when to use them.

List all of the strategies you will use to identify and manage bots or imposter participants in your research ethics proposal. Ensure all screening and verification methods comply with **ethical guidelines**, respect participant privacy, and are approved by the Research Ethics Board at the University of Alberta.

Throughout the research process, **maintain detailed records** of your processes, instances of suspected fraudulence, and rationale for the exclusion of a participant or the data collected.

Summary

By implementing these strategies and remaining vigilant throughout your research process, you can significantly reduce the risk of bots and imposter participants compromising the integrity and validity of your valuable research.

Again, use your judgment and consider your study design, the unique needs of your study population, and sampling methods when deciding which strategies to use.

Remember to consult with the Research Ethics Board for guidance on the most appropriate and ethical methods for your specific study.

Additional Resources

Brooks, H.M., Erdmann, G., Scott, S.D. (2025). Digital disguises: Examining the impacts of imposter participants in online research. Poster. 2025 NorthWest SPOR Collaborative Forum. Adapted poster available online here.

Fernandes, A., Macdonald, K., Pardo, M. C., Jansen, M., Page, S., & Zelinsky, S. (2023, October 4). 15 fake participants in online patient-oriented research: How it looks and how to navigate it [Webinar recording]. Alberta SPOR Support Unit (AbSPORU). https://vimeo.com/876417825

Pullen Sansfaçon A, Gravel E, Gelly MA. Dealing with scam in online qualitative research: strategies and ethical considerations. Int J Qualitative Methods. 2024;23:16094069231224610. https://doi.org/10.1177/16094069231224610

Roehl JM, Harland DJ. Imposter participants: overcoming methodological challenges related to balancing participant privacy with data quality when using online recruitment and data collection. Qualitative Rep. 2022;27(11):2469–85. https://doi.org/10.46743/2160-3715/2022.5475

Ridge D, Bullock L, Causer H, Fisher T, Hider S, Kingstone T, et al. Imposter participants' in online qualitative research, a new and increasing threat to data integrity? Health Expectations: Int J Public Participation Health Care Health Policy. 2023;26(3):941. https://doi.org/10.1111/hex.13724

Santinele Martino, A., Perrotta, A., & McGillion, B. J. (2024). Who can you trust these days?: Dealing with imposter participants during online recruitment and data collection. Qualitative Research, 24(5), 1291-1301. https://doi.org/10.1177/14687941231224591

Sharma, P., McPhail, S.M., Kularatna, S. et al. (2024). Navigating the challenges of imposter participants in online qualitative research: lessons learned from a paediatric health services study. BMC Health Serv Res 24, 724 https://doi.org/10.1186/s12913-024-11166-x

University of Alberta ARISE example. Available online here.

This resource was developed collaboratively by the Faculty of Nursing and the School of Public Health at the University of Alberta.