Exploring Scaffolding Techniques for Agent-Administered Brief Cognitive Screening in Hospital Settings

Dingdong Liu
The Hong Kong University of Science and Technology
Hong Kong, China
dliuak@connect.ust.hk

Sensen Gao
Nankai University
Tianjin, China
gaosensen@mail.nankai.edu.cn

Zixin Chen
The Hong Kong University of Science and Technology
Hong Kong, China
zchendf@connect.ust.hk

Yifan Shen
The Hong Kong University of Science and Technology
Hong Kong, China
yshenaz@conncet.ust.hk

Chuhan Shi
Southeast University
Nanjing, China
chuhanshi@seu.edu.cn

Bertram E. Shi
The Hong Kong University of Science and Technology
Hong Kong, China
eebert@ust.hk

Xiaojuan Ma*
The Hong Kong University of Science and Technology
Hong Kong, China
mxj@cse.ust.hk

ABSTRACT

Cognitive screening in hospitalized older patients is a critical, yet time-consuming process. While conversational agents present a promising solution to aid clinicians, current models fall short in their ability to scaffold questions to accommodate patients with potential cognitive decline effectively. To bridge this gap, we conducted a study with 13 clinicians to identify effective scaffolding strategies empirically. Our findings revealed six key strategies that clinicians use to scaffold the Abbreviated Mental Test (AMT) in practice, together with the underlying rationale and potential challenges. We discuss the implications of these findings for the design of conversational agents to assist in cognitive screening and propose design considerations for future research.

CCS CONCEPTS

• **Human-centered computing** → **Participatory design**; **Empirical studies in interaction design**; **Empirical studies in HCI**; **Natural language interfaces**.

KEYWORDS

conversational agent, scaffolding, cognitive screening

*Corresponding Author

DIS Companion '24, July 1–5, 2024, IT University of Copenhagen, Denmark

ACM ISBN 979-8-4007-0632-5/24/07…$15.00
https://doi.org/10.1145/3656156.3663697

ACM Reference Format:
Dingdong Liu, Sensen Gao, Zixin Chen, Yifan Shen, Chuhan Shi, Bertram E. Shi, and Xiaojuan Ma. 2024. Exploring Scaffolding Techniques for Agent-Administered Brief Cognitive Screening in Hospital Settings. In *Designing Interactive Systems Conference (DIS Companion '24), July 1–5, 2024, IT University of Copenhagen, Denmark.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3656156.3663697

1 INTRODUCTION

Cognitive screening, a rapid evaluation to identify individuals needing comprehensive neuropsychological testing, is vital in clinical practice for older adults and those with potential cognitive impairments [2, 22]. These tests are regularly administered to each patient in need. Several oral tests, including the Abbreviated Mental Test (AMT) [5, 9], Mini-Mental State Examination (MMSE) [6], Telephone Interview for Cognitive Status (TICS) [3], and Telephone Montreal Cognitive Assessment (T-MoCA) [16], are commonly used. These tests, requiring clinician administration, typically take 5 to 10 minutes to complete.

While beneficial, regular cognitive screening can considerably burden clinicians. In East Asia's public hospitals, clinicians often screen over 20 patients daily, consuming substantial time. This workload can cause fatigue and burnout, potentially diminishing care quality. Moreover, screening frequency may drop during peak times, leading to potential diagnostic and intervention delays that can exacerbate prolonged cognitive impairments [8]. As conversational agents have the advantage of consistency, scalability, and cost-effectiveness, there is a growing interest in involving conversational agents in medical services [11, 19, 23]. However, they cannot be directly applied to the cognitive screening process due to the unique support needed by the patients taking cognitive screening. In the screening process, clinicians must ensure that patients comprehend the questions and provide adequate information per the assessment guidelines. This may be challenging when patients

undergoing the assessment experience varying degrees of cognitive decline, which may affect expressive and receptive communication functions, including memory, attention, executive function, visual-spatial skills, and language [10, 17, 20]. These patients might require additional communication support, such as metaphors or analogies to comprehend the questions, reminders to track the questions, and appropriate assistance to maintain their autonomy while avoiding frustration due to errors [18, 20]. In such situations, scaffolding, a common instruction technique that prompts people to elicit information through questions and hints, is widely adopted by clinicians [18].

Nevertheless, previous research on scaffolding has primarily focused on children's and adults' learning processes, not on older adults' cognitive screening. Besides the difference in scenarios, it remains unclear what scaffolding strategies can effectively guide patients without compromising the test's integrity. For example, simplifying the question might be a more preferred strategy than directly providing the answer. Similar scaffolding processes are observed in other healthcare services related to patient interviews, such as telephone triage, recovery process tracking, and post-stroke rehabilitation. Celino et al. [4] and Levin et al. [13] provided static reflective questions and examples to assist patient's question understanding and answering. However, they did not provide design guidelines for these questions, and patients criticized the provided pre-defined questions for not being specific to their situations.

To address the challenges above, our research aims to investigate the scaffolding strategies employed by clinicians during cognitive screening. In this qualitative study, we engaged 13 clinicians from various occupations, all with extensive experience in cognitive screening. We first presented a prototype of a cognitive screening conversational agent to the clinicians to demonstrate the potential of using conversational agents to support clinicians in this process. Subsequently, we conducted semi-structured interviews to gain insights into the empirical scaffolding strategies adopted by clinicians during cognitive screening and their recommendations for the design of conversational agents in this context. Through analysis of the interview data, we identified six key scaffolding strategies employed by clinicians during cognitive screening, along with their recommendations for designing conversational agents for cognitive screening. Our pioneering work uncovers empirical skills used in cognitive screening, providing design requirements for future conversational agents. By identifying effective scaffolding strategies and their integration into these agents, we present a potential tool to enhance cognitive assessment efficiency, bridging the gap between technology and healthcare.

2 METHODS

To gain insights into how clinicians support the patient in cognitive screening, particularly the scaffolding strategies clinicians employ, we conducted a study involving 13 clinicians from a hospital in an East Asian city. The study investigates the scaffolding strategies employed by clinicians during the AMT test and the reasoning behind them through a role-play session and a semi-structured interview.

2.1 Prototype System

We developed a prototype conversational agent running on a humanoid robot[1], that delivered a localized version of the Abbreviated Mental Test (AMT) [9], a widely used cognitive screening test worldwide, including the local hospital we worked with. A prototype system was implemented with the functions following preview studies in patient interviews, including cognitive tests [21], diabetes self-management [15], and patient information collection [1, 7]. These functions include providing welcoming messages, raising questions according to a preset script, responding to user's answers, and expressing gratitude at the end of the test. The conversation between the user and robot was controlled by a dialogue manager implemented using the Dialogflow[2]. Once the user responds to a question, the robot will transcribe[3] the response and send it to the dialogue manager. Slot-filling against the question is performed. The dialogue manager then decides to either (1) repeat the latest question, (2) move to the next question, or (3) disengage the dialogue when the test is completed. Corresponding utterances are selected from the preset script by the dialogue manager and then sent to the voice module for synthesis. The robot's voice module[4] will then play the synthesized speech to the user.

Furthermore, we have integrated functionality to detect the user's attention and emotional state using the built in voice activity detection (VAD) and facial expression recognition algorithms. This enables us to identify instances when the user becomes distracted[5] and allows for the repetition of questions if necessary. Additionally, in cases where the user becomes agitated[6], the system can gracefully disengage and conclude the test.

2.2 Participants and Procedure

With the approval of the institution's IRB, we recruited 13 clinicians from a local hospital through the hospital's internal communication channels. Demographic information about the participants can be found in Table 1.

Table 1: Participants' demographic information. In this table, "Title" refers to the participant's job title. APN stands for Advanced Practice Nurse, RN stands for Registered Nurse, OT stands for Occupational Therapist, and WM stands for Ward Manager.

ID	Gender	Age	Title	ID	Gender	Age	Title
P1	Male	25-34	APN	P8	Female	25-34	RN
P2	Female	35-44	RN	P9	Female	25-34	APN
P3	Male	25-34	RN	P10	Female	25-34	RN
P4	Male	35-44	APN	P11	Female	35-44	APN
P5	Female	45-54	APN	P12	Female	35-44	OT
P6	Male	45-54	APN	P13	Female	25-34	OT
P7	Female	45-54	WM				

[1]https://awakening.health
[2]Implemented with Google's Dialogflow: https://cloud.google.com/dialogflow
[3]ASR is provided by Google Cloud: https://cloud.google.com/speech-to-text?hl=en
[4]Implemented with Amazon Polly: https://aws.amazon.com/polly/
[5]When users do not answer a question (no human voice) for more than 1.5 seconds, they are classified as distracted.
[6]When users' facial expressions are classified as anger or disgust, they are perceived as agitated.

All participants had previous experience administering the AMT test to elderly patients, although they held varying job titles. None of them had prior experience with conversational agents, except for P13, who had experience in using a chatbot for patient information collection. According to feedback from clinicians before the experiment, the AMT tests are conducted daily in hospital rooms where patients typically lie on a bed. To replicate this environment, we created a simulated setting for our experiment. During the experiment, the robot was placed at the bedside, and the participants were asked to interact with the robot as if they were the patient. We requested the participants to role-play the patients they would typically encounter in their daily work and to respond to the robot's questions as they would in a real-life scenario. Each participant was given the opportunity to role-play four times, each time trying different types of patients, resembling the cases they encounter in their daily work. Before the role-play, we knew from the participants that some patients needed assistance with the test. Therefore, we asked the participants to role-play twice as patients who may need assistance in the test. The order of the role-playing as a patient needing assistance was randomly distributed in the four role-plays to mitigate any potential order effects. The role-play sessions were video-recorded for subsequent interview and analysis.

Following the role-play sessions, each participant was invited to a semi-structured interview to share their insights and experiences. The interview mainly focuses on two parts: (1) **Experience when giving the AMT test in the hospital.** Questions include how they conducted the test, the frequency of the test, the challenges they encountered, and the methods they used to overcome the challenges. (2) **A retrospective think-aloud process.** Participants were invited to review the video of their role-plays, explain the characteristics of the patient they want to replicate, and comment on the robot's behavior. The interview was conducted in a separate room to avoid mutual influence, and it was recorded and transcribed for analysis.

2.3 Analysis

Two researchers used deductive and inductive approaches to analyze the interview transcripts thematically. Two researchers first independently coded four transcripts related to the robot's behavior. These preliminary codes and themes were then discussed and agreed upon. The researchers independently coded the remaining transcripts, ultimately reaching a consensus on the final themes and codes. The findings include six identified empirical scaffolding strategies adopted by the nurses (Table 2). For each scaffolding strategy, the researchers revisited the transcripts to gather sample responses. Krippendorff's alpha was used to compute the inter-rater reliability (IRR) to ensure the quality of sample response collection, as recommended by McDonald [14]. The IRR reached 0.895, denoting an acceptable level of agreement [12]. The two researchers then discussed the discrepancies and agreed on the final labels. Table 2 shows some examples of the identified scaffolding strategies.

3 RESULTS

Overall, the clinicians positively approached using conversational agents in cognitive screening. All but P4 believed the humanoid robot feels warm and friendly and may help reduce patients' anxiety. While they thought the current system works with many patients, some places still need improvement, especially regarding the robot's ability to scaffold the questions and maintain the test flow. We outline the key findings from the study below.

Purpose of scaffolding in AMT test. According to the clinicians, the AMT test is designed to be standardized, and clinicians are advised against altering the wording of the questions. However, it is common for patients to have difficulty understanding the questions or providing fully correct answers on the first attempt. As a result, clinicians often employ various techniques to scaffold the questions and support the patients. For example, *"I will give some kind of not hint, but to guide the patients to tell me what they know. (P13)"*. Additionally, participants noted that some patients might respond with "I don't know" to every question. Under such circumstances, the purpose of scaffolding is to encourage the patient to try their best (P12).

Empirical scaffolding strategies adopted by clinicians. We identified six approaches to scaffolding through the participants' feedback, categorized as "follow-up question", "give hints", "ask to think again", "ask for a guess", "change wording", and "simplify question". The participants mentioned that they will try to ask follow-up or further questions, give hints, or ask the patient to think again when they give an approximate or partially correct answer. E.g., *"only give the year but not the month when asked about the date (P7)"*, *"the patient may only reply in the morning when asked about time (P10)"*. When the patient cannot answer, the clinicians will consider if the patient understands the question and will try to change the wording. Besides, some questions in the AMT test accept multiple answers, and the clinicians will try to ask the patient to give a simplified answer, such as *"I will ask in daytime, afternoon or at night (P13)"*. In addition, the clinicians will also ask the patient to make a guess when the patient cannot answer after several tries. *"Though the patient may not be able to give a correct answer, we still have a clue about the patient's cognitive status"*, as is mentioned by P13.

Maintain the flow of the test. Although scaffolding can be useful during testing, it may also prolong the duration of the test and frustrate the patient. Clinicians have emphasized the importance of maintaining the test flow, as sticking to a question for too long can be detrimental. Patients may lose their attention and confidence or even refuse to cooperate if the test lasts too long. To prevent this from happening, clinicians suggest that the robot can give quick and encouraging responses such as "good job" or "you speak so well" which can effectively motivate the patients to continue with the test. (P1, P2, P3). Participants also stressed the importance of confirming that patients follow the robot's instructions. For instance, P12 mentioned instances where patients may still be processing the previous question while the AMT administrator has already moved on to the next one. *"I will ask the question again to make sure she can hear clearly and focus on the new questions, but not occupied by the old one. (P12)"* On the other hand, some clinicians mentioned that they would avoid sticking to a question by temporarily omitting some questions. *"In tests that sequence doesn't matter, I will skip the question [that patient cannot answer] and go back to it later (P7)"*. Sometimes, the clinicians also encounter patients who lose patience and refuse to cooperate. This time, the clinicians suggest the robot

Table 2: Scaffolding strategies used by clinicians during cognitive screening, along with the number of identified scenarios and frequency of assistance needed in patient role-plays. The frequency is divided into two cases: case 1 refers to role-plays where clinicians were explicitly required to role-play as patients that may need assistance prior to their role-play, while case 2 role-plays did not have such an explicit requirement.

Scaffolding Strategies	Examples	Scenarios	Frequency	
			Case 1	Case 2
Follow-up question	(when the patient only gives the time) need to check further if it is AM/PM	11	5/26	9/26
Give hints	"Look outside the window, is it night or morning now?"	7	7/26	1/26
Ask to think again	"If she said she is in her 50s (which is wrong), then I ask her to think again."	5	7/26	2/26
Ask for a guess	(after several failures or "I don't know") "Then I'll ask him/her to try to guess."	5	11/26	0/26
Change wording	"Change the wording when repeating questions (to ensure understanding)"	2	8/26	0/26
Simplify question	I will ask in approximate - are you in your 80s or 60s - if patient cannot remember exact age.	2	5/26	2/26

disengages the conversation and reports to the nurse (P2, P10, P11, P13).

4 LESSONS LEARNED AND FUTURE WORKS

In conclusion, we summarize three main lessons learned from the study that will inform our future research and development of conversational agents for cognitive screening.

Support patients with scaffolding strategies. We discovered the empirical scaffolding strategies used by clinicians during the AMT test, including "follow-up question", "giving hints", "asking to think again", "asking for a guess", "change wording", and "simplify question". Scaffolding is necessary as it encourages patients to express more about themselves, so that the clinicians can better understand their medical needs. Future work could consider incorporating these strategies into the conversational agent to support patients during the test. According to our empirical findings, the most common pattern of scaffolding involves first assessing the correctness of a patient's answers, and then selecting a scaffolding strategy based on the necessary level of assistance. Future designs could benefit from mimicking such a decision-making process, as informed by our empirical findings.

Identify the appropriate degree of scaffolding. While patients may need assistance during the test, clinicians also emphasized the importance of maintaining the test flow and suggested various approaches to achieve this goal. Scaffolding can aid patients during cognitive screening, but excessive scaffolding may distract or even annoy them. This suggests that the designed conversational agent should consider the patient's emotional status and adjust the degree of scaffolding accordingly. Future work needs to balance assisting the patient with maintaining the test flow when designing the overall conversational workflow.

Incorporating actual patients. The research was conducted in a simulated environment, with domain experts serving as proxies for patients in various cognitive states. This approach was necessary due to the inherent challenges in recruiting the target population, particularly individuals with cognitive decline, given their medical conditions. Additionally, the clinicians we engaged have significant experience in administering the AMT test and possess practical field experience, equipping them to comprehend the reactions and requirements of the patients. Nonetheless, future work should consider including actual patients to validate findings and gain deeper insights into their genuine challenges and needs. This could take the form of an iterative co-design study, gradually integrating more clinicians and real patients into the development stages. Direct feedback from stakeholders could ensure that the conversational agents would effectively improve engagement and outcomes during cognitive screenings. Future studies might also examine the long-term impact of such agents in clinical practices, providing more profound insights into their practical benefits and limitations.

ACKNOWLEDGMENTS

Many thanks to the anonymous reviewers for their insightful suggestions. This work is supported by the Hong Kong University of Science and Technology (Research Program Development in Aging, Healthcare and Mental Wellbeing) with award number: Z1011/FK081.

REFERENCES

[1] Ho Seok Ahn, Wesley Yep, Jongyoon Lim, Byeong Kyu Ahn, Deborah L. Johanson, Eui Jun Hwang, Min Ho Lee, Elizabeth Broadbent, and Bruce A. Macdonald. 2019. Hospital Receptionist Robot v2: Design for Enhancing Verbal Interaction with Social Skills. *2019 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019* (2019), 1–6. https://doi.org/10.1109/RO-MAN46459.2019.8956300

[2] Patrícia Belchior, Nicol Korner-Bitensky, Melanie Holmes, and Alexandra Robert. 2015. Identification and Assessment of Functional Performance in Mild Cognitive Impairment: A Survey of Occupational Therapy Practices. *Australian Occupational Therapy Journal* 62, 3 (2015), 187–196. https://doi.org/10.1111/1440-1630.12201

[3] Jason Brandt, Miriam Spencer, and Marshal Folstein. 1988. The Telephone Interview for Cognitive Status. *Cognitive and Behavioral Neurology* 1 (Jan. 1988).

[4] Irene Celino and Gloria Re Calegari. 2020. Submitting Surveys via a Conversational Interface: An Evaluation of User Acceptance and Approach Effectiveness. *International Journal of Human-Computer Studies* 139 (July 2020), 102410. https://doi.org/10.1016/j.ijhcs.2020.102410

[5] LW Chu, CKW Pei, MH Ho, and PT Chan. 1995. Validation of the abbreviated mental test (Hong Kong version) in the elderly medical patient.

[6] M. F. Folstein, S. E. Folstein, and P. R. McHugh. 1975. "Mini-mental State". A Practical Method for Grading the Cognitive State of Patients for the Clinician. *Journal of Psychiatric Research* 12, 3 (Nov. 1975), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

[7] Eui Jun Hwang, Bruce A. MacDonald, and Ho Seok Ahn. 2019. End-to-End Dialogue System with Multi Languages for Hospital Receptionist Robot. *2019 16th International Conference on Ubiquitous Robots, UR 2019* (2019), 278–283. https://doi.org/10.1109/URAI.2019.8768694

[8] Catherine E. Jansen, Bruce A. Cooper, Marylin J. Dodd, and Christine A. Miaskowski. 2011. A Prospective Longitudinal Study of Chemotherapy-Induced Cognitive Changes in Breast Cancer Patients. *Supportive Care in Cancer* 19, 10 (Oct. 2011), 1647–1656. https://doi.org/10.1007/s00520-010-0997-4

[9] S. Jitapunkul, I. Pillay, and S. Ebrahim. 1991. The Abbreviated Mental Test: Its Use and Validity. *Age and Ageing* 20, 5 (Sept. 1991), 332–336. https://doi.org/10.1093/ageing/20.5.332

[10] Melissa Johnson and Feng Lin. January/March 2014. Communication Difficulty and Relevant Interventions in Mild Cognitive Impairment: Implications for Neuroplasticity. *Topics in Geriatric Rehabilitation* 30, 1 (January/March 2014), 18. https://doi.org/10.1097/TGR.0000000000000001

[11] Junhan Kim, Sun Park, and Lionel + "Jr" Robert. 2019. Conversational Agents for Health and Wellbeing: Review and Future Agendas. CSCW 2019.

[12] Klaus Krippendorff. 2004. Reliability in Content Analysis: Some Common Misconceptions and Recommendations. *Human Communication Research* 30, 242 (July 2004), 411–433.

[13] Esther Levin and Alex Levin. 2006. Evaluation of Spoken Dialogue Technology for Real-Time Health Data Collection. *Journal of Medical Internet Research* 8, 4 (Dec. 2006), e30. https://doi.org/10.2196/jmir.8.4.e30

[14] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and Inter-rater Reliability in Qualitative Research: Norms and Guidelines for CSCW and HCI Practice. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–23. https://doi.org/10.1145/3359174

[15] Mark A. Neerincx, Willeke Van Vught, Olivier Blanson Henkemans, Elettra Oleari, Joost Broekens, Rifca Peters, Frank Kaptein, Yiannis Demiris, Bernd Kiefer, Diego Fumagalli, and Bert Bierman. 2019. Socio-cognitive engineering of a robotic partner for child's diabetes self-management. *Frontiers in Robotics and AI* 6, November (2019). https://doi.org/10.3389/frobt.2019.00118

[16] Sarah T. Pendlebury, Sarah J.V. Welch, Fiona C. Cuthbertson, Jose Mariz, Ziyah Mehta, and Peter M. Rothwell. 2013. Telephone Assessment of Cognition After Transient Ischemic Attack and Stroke. *Stroke* 44, 1 (Jan. 2013), 227–229. https://doi.org/10.1161/STROKEAHA.112.673384

[17] R. C. Petersen. 2004. Mild Cognitive Impairment as a Diagnostic Entity. *Journal of Internal Medicine* 256, 3 (Sept. 2004), 183–194. https://doi.org/10.1111/j.1365-2796.2004.01388.x

[18] Pamela A. Saunders, Kate de Medeiros, and Ashley Bartell. 2011. 'Oh He Was Forgettable': Construction of Self Identity through Use of Communicative Coping Behaviors in the Discourse of Persons with Cognitive Impairment. *Dementia* 10, 3 (Aug. 2011), 341–359. https://doi.org/10.1177/1471301211407805

[19] Korok Sengupta, Sayan Sarcar, Alisha Pradhan, Roisin McNaney, Sergio Sayago, Debaleena Chattopadhyay, and Anirudha Joshi. 2020. Challenges and Opportunities of Leveraging Intelligent Conversational Assistant to Improve the Well-being of Older Adults. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–4. https://doi.org/10.1145/3334480.3381057

[20] Claire Tochel, Michael Smith, Helen Baldwin, Anders Gustavsson, Amanda Ly, Christin Bexelius, Mia Nelson, Christophe Bintener, Enrico Fantoni, Josep Garre-Olmo, Olin Janssen, Christoph Jindra, Isabella F. Jørgensen, Alex McKeown, Buket Öztürk, Anna Ponjoan, Michele H. Potashman, Catherine Reed, Emilse Roncancio-Diaz, Stephanie Vos, and Cathie Sudlow. 2019. What Outcomes Are Important to Patients with Mild Cognitive Impairment or Alzheimer's Disease, Their Caregivers, and Health-Care Professionals? A Systematic Review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 11 (Dec. 2019), 231–247. https://doi.org/10.1016/j.dadm.2018.12.003

[21] Simone Varrasi, Santo Di Nuovo, Daniela Conti, and Alessandro Di Nuovo. 2018. A Social Robot for Cognitive Assessment. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18)*. Association for Computing Machinery, New York, NY, USA, 269–270. https://doi.org/10.1145/3173386.3176995

[22] H.J. Woodford and J. George. 2007. Cognitive Assessment in the Elderly: A Review of Clinical Methods. *QJM: An International Journal of Medicine* 100, 8 (Aug. 2007), 469–484. https://doi.org/10.1093/qjmed/hcm051

[23] Tamara Zubatiy, Kayci L Vickers, Niharika Mathur, and Elizabeth D Mynatt. 2021. Empowering Dyads of Older Adults With Mild Cognitive Impairment And Their Care Partners Using Conversational Agents. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3411764.3445124