When to Explain: Field Study Insights on Robot Failure Explanations for Older Adults

Tamlin Love*
tlove@iri.upc.edu
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Llorens i Artigas 4-6, 08028
Barcelona Spain

Víctor Bermejo*
victor.bermejo@iiia.csic.es
Artificial Intelligence Research Institute (IIIA-CSIC), Universitat Autònoma de Barcelona (UAB)
08193 Bellaterra, Barcelona Spain

Alberto Olivares-Alarcos
aolivares@iri.upc.edu
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Llorens i Artigas 4-6, 08028
Barcelona Spain

Antonio Andriella
aandriella@iri.upc.edu
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Llorens i Artigas 4-6, 08028
Barcelona Spain

Núria Vallès-Peris
nuria.valles@iiia.csic.es
Artificial Intelligence Research Institute (IIIA-CSIC), Universitat Autònoma de Barcelona (UAB)
08193 Bellaterra, Barcelona Spain

Cristian Barrué
Guillem Alenyà
cbarrue@iri.upc.edu
galenya@iri.upc.edu
Institut de Robòtica i Informàtica Industrial (CSIC-UPC)
Llorens i Artigas 4-6, 08028
Barcelona Spain

Abstract

Explainability has been proposed as an approach to robot failure recovery, facilitating understanding and repairing trust, especially relevant in domestic assistive tasks. This study conducts a preliminary exploration of older adults' preferences regarding the content and context of robot-generated explanations for failures to guide future research. An exploratory study was conducted in three phases: 1) gathering high-level requirements from caregivers, 2) implementing a semi-autonomous robot for object retrieval that identifies and explains different types of failures, and 3) engaging N=8 older adults in real-life interactions as well as in focus groups to assess their perspectives. Our preliminary observations highlight a tension in preferences: a general desire for short, direct explanations to minimize disruption, versus a need for more detailed, actionable explanations specifically in failure cases. Crucially, we also note that these preferences are unstable and contextually constructed, reinforcing that the technical failures cannot be separated from their social context, as users' experiences and opinions are shaped by both the robot's functional capabilities and the values and organisational settings in which they are introduced.

CCS Concepts

• **Human-centered computing → Interaction design process and methods**; **Field studies**.

Keywords

Explainability, Failure Explanations, Human-Robot Interaction

*Both authors contributed equally to this research.

HRI Companion '26, Edinburgh, Scotland, UK

ACM ISBN 979-8-4007-2321-6/2026/03
https://doi.org/10.1145/3776734.3794462

ACM Reference Format:
Tamlin Love, Víctor Bermejo, Alberto Olivares-Alarcos, Antonio Andriella, Núria Vallès-Peris, Cristian Barrué, and Guillem Alenyà. 2026. When to Explain: Field Study Insights on Robot Failure Explanations for Older Adults. In *Companion Proceedings of the 21st ACM/IEEE International Conference on Human-Robot Interaction (HRI Companion '26), March 16–19, 2026, Edinburgh, Scotland, UK*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3776734.3794462

1 Introduction

As populations age, there is an increasing focus on "ageing in place", where older adults maintain their independence by continuing to live in their own homes rather than transitioning to assisted living facilities [4]. Assistive technologies, such as home robots, have emerged as a promising solution to support older adults in their daily activities, providing physical assistance [19] and cognitive and social stimulation [3, 15]. However, due to the technical limitations on robots' capabilities and the complexity of domestic assistive tasks, it is likely that such robots will experience failures of varying kinds, which can negatively impact trust and acceptance [7].

One strategy for recovering from such failures is for the robot to produce explanations for why the failure occurred [21]. If a robot can explain why it failed to complete a task, it can foster an improved understanding of the robot's decision-making in the user, and bridge the gap between the user's expectations of the robot and the reality of its capabilities and reasoning [7]. Indeed, work in this area has highlighted many important factors impacting a failure explanation's impact, including the timing [13], inclusion of context [8], type of failure and the user's situational awareness [12].

To successfully deploy robots in users' homes, it is crucial to include end users, such as older adults and caregivers, early in the design process, so that both the functionality and explanations of a robot are relevant and useful for end users [11]. Such an inclusive approach is especially important for older adults, whose experiences and perspectives are often overlooked in the design of technologies [14]. While there has been some work co-designing

(a) Co-Design with Caregivers

(b) Implementation of Task

(c) In Situ Session

(d) Focus Group

Figure 1: An overview of our study: (a) initial requirements gathering sessions with caregivers; (b) the implementation of the object retrieval task and failure explanations; (c) in situ sessions with participants to familiarise them with the robot and gather feedback and (d) focus group sessions to gather additional feedback and brainstorm improvements for the robot's explanations.

robotic solutions for older adults in domestic environments [15, 18], to the best of our knowledge, this work represents one of the first focused explorations into failure explanations for assistive tasks for older adults in domestic environments.

Therefore, this study aims to explore the following research question: *What failure explanations are preferred by older adults when it comes to understanding and interacting with a household robot?* To that end, we conduct a qualitative study consisting of four phases (see Fig. 1): 1) requirements gathering with caregivers, 2) the implementation of an object retrieval task for a robot in domestic environments, including the identification and explanation of failures, 3) one-on-one sessions with older adults interacting with the robot in their own homes, and 4) focus group sessions with participants to gather additional feedback and reflections. This study suggests that participants prefer short, direct explanations only when necessary, prioritising task performance and minimising disruption to daily life. We also provide a number of insights on the contextual impacts of the domestic environment on the relationships between older adults and the robot, which may prove useful for future field studies in explainable, assistive robotics.

2 Initial Design

To better align our study with the needs of older adults in their homes, we began by conducting workshops and one-on-one interviews with caregivers[1]. The aim of these initial sessions was to identify tasks for which robotic intervention could meaningfully assist older adults and how explainability could factor into the task design. Together with the caregivers, we identified object retrieval, where a robot could assist users in finding and retrieving objects around the house, as a useful task. As navigation, object detection and grasping are all challenging, especially in domestic contexts, we noted that failure is likely to occur, highlighting the need to identify and explain failures. Following these requirements, we implemented the object retrieval task[2] for the Stretch RE2 (see Fig. 1b), which has seen use in a number of indoor assistive robot studies [2, 17]. Upon receiving the command to fetch an object, the robot searches each room in a house until the object is found. It then grasps the object (if found) and returns to the user to hand it over.

We identify two particularly relevant failure types for our in situ study: *navigation failure*, which occurs if the robot cannot reach a destination, and *grasping failure*, which occurs if the robot is not able to grasp an object. Given these possible failures, we define a number of templated natural language explanations that can be provided at various stages of the task. The first is the *narrative explanation*, which describes a given robot execution. It can be given at the start of a task, describing a planned trajectory (e.g., "I will go to the kitchen to find the cup.") or at the end of a task, describing a past trajectory (e.g., "I went to the kitchen and then tried to go to the bathroom, but my path was blocked."). The second type is the *expectation briefing*, where the robot informs the user about possible attributes of its future task. This could set expectations of possible failures (e.g., "It's possible I could fail: I might not find the cup, or my way could be blocked, or I might struggle to pick it up."), or about other attributes such as the time to complete the task.

The robot can also provide *task confirmation* utterances when a user gives a command (e.g., "I will fetch you the cup.") and *task outcome* utterances to express whether the task succeeded or failed (e.g., "Here is the cup" in a success case, "I wasn't able to bring you the cup" in a failure case). Finally, a *progress update* combines elements of a narrative explanation and a task outcome to provide updates on the robot's task progress during its execution (e.g., "I haven't found the cup yet, I will try the dining room next.").

3 Field Study with Older Adults

To assess how robot explanations are perceived in a domestic context, we conducted a study in which $N = 8$ older adults (5 male, 3 female, age $M = 82.38$, $STD = 7.25$) interacted with the robot in their own homes[3]. The participants were recruited following a purposive sampling of volunteers from residents of two sheltered houses for independent older adults, ensuring variability in their gender and ages. Each session was conducted in either Catalan or Spanish. During individual, in situ sessions, we conducted semi-structured interviews that allowed us to gather standardised information about participants while exploring in a deeper way other ideas that may emerge over the course of the interview. We then conducted two focus group sessions (FG1 $N = 3$; FG2 $N = 3$) in which participants came together to reflect and provide more detailed feedback about their experience with the robot. Two participants (P3 and P6) abstained from participating for personal reasons. Both interviews and focus groups were analysed using a thematic analysis. We coded participants' utterances concerning robot explanations (length, frequency and type), task performance (attitudes towards success or failure), and interactional experience (volume, space navigation, relevance of the tasks, enjoyment, boredom).

[1] https://www.suara.coop/en

[2] Details may be found at https://github.com/tamlinlove/domestic_stretch

[3] The study was approved by the ethics committee of the Spanish National Research Council (CSIC), reference number 228/2024.

The purpose of the individual in situ sessions was to familiarise participants with the robot and the task, especially regarding failure and corresponding explanations, and to gather feedback about the experience in their homes. We conducted this session prior to more free-form focus group sessions in order to ground future brainstorming sessions in the reality of the robot's embodiment and implementation. This is especially important in the case of older adults, whose familiarity with robots may be low [1].

Each session followed a fixed sequence. First, the interviewer, a PhD candidate with a background in social sciences and expertise in qualitative analysis, introduced the participant (*Pi*) to the session goals, described the robot's capabilities, and set expectations, emphasising that the robot might fail and that the session aimed to explore how robots should explain such failures. The interviewer also asked brief questions to assess the participant's socio-economic background and familiarity with robots and technology. After this introduction, the participant completed five object-retrieval trials (T*i*), each with a different configuration of failures, explanations and discussion topics, as outlined in Table 1. In each trial, the participant requested an object and the trial ended when the robot returned, with or without the object, depending on failures. Discussions between the interviewer and participant occurred before and after the task execution. Afterwards, we conducted a debriefing session to explain the study's aims and answer any questions.

4 Analysis of Findings

4.1 In Situ Sessions

Participants tended to prefer short, simple and direct explanations. While, at first, the fact that the robot could express itself was entertaining, their attention soon shifted to whether the robot could perform the task and navigate the home environment. Participants expressed a desire for the robot to perform the task efficiently, quickly and unobtrusively, and this extended to its explanations.

P2: *Just let it go for the bottle and don't overwhelm me any further.*

However, for T3 and T4, in which failures occurred, participants showed interest in the robot providing more in-depth explanations about the causes of failures, allowing them to take action to ensure successful task completion in the future. This matches findings in the literature, suggesting explanations should be actionable [6] and should only be given when the unexpected occurs [21].

P7: *What I'd like is that if I tell the robot to grab something and it can't, for whatever reason, it should let me know: "I wanted to grab the orange, but there was an obstacle". Then you have to go there and see what the obstacle is.*

While we attempted to assess participants' mental models of the robot's decision-making in T5, participants were not eager to take the robot's perspective. Instead, some users showed interest in providing instructions to the robot on where to find the requested objects. They were not particularly interested in the robot's mental model itself, but rather in its ability to easily translate their own particular understanding of their household to the robot by giving it instructions. Participants argued that they know their home and where they store their belongings, so it would be more efficient for them to guide the robot through the living space, facilitating task completion and reducing the time required.

P3: *But it would be useful if you could tell it where to search. Because earlier, I told it "Bring the remote," and it started looking for it. But if I say "Bring the remote that's in the bedroom," would that make it faster?*

Participants did not show a particular preference regarding explanation types, they were more concerned with length and frequency. They preferred a robot to be as unobtrusive as possible in both behaviour and explanations. They wanted the robot to integrate seamlessly into their daily routines, which could be interrupted by explanations. Thus, shorter, sparser and more direct explanations are less disruptive. Although some users highly valued having conversations with the robot, their preferences for explanations suggest that they favoured verbal interaction focused on specific tasks and moments rather than sustained communication throughout their execution of daily routines.

4.2 Focus Group Sessions

During the focus group sessions, we were surprised by two factors: (1) the two focus groups displayed opposite preferences regarding the robot's explanations and behaviour, and (2) these preferences seemed to have changed significantly for some participants compared to their initial reactions during the in situ sessions. Contrary to our initial findings, the users in *FG1* were in favour of longer explanations. For them, the robot talking about its tasks and providing information contributed to a sense of companionship.

P2: [*FG1*] *I think so, I did like it [...] Because, to begin with, you have company.*

However, the users in the *FG2* preferred explanations to be as short as possible and, if feasible, would rather not have them at all. Even when asked about explanations in the case of failure, they maintained their preference for their absence. For them, explanations delayed task completion and disrupted the flow of interaction.

P4: [*FG2*] *For the robot to be useful to me [...] it has to listen to me and do it. Not explain what it's going to do.*

P7: [...] *If I must have a robot and wait 5 minutes for it to explain something to me, or if I have to explain something to it, some people will get very nervous.*

5 Discussion

No stable or universal preference emerged regarding how robots should explain failures. Instead, preferences were deeply context-dependent. The same people have different preferences depending on the context in which the question was asked, either at home during the in situ session or retrospectively during the focus groups. However, this does not imply that the answers are irrational or irrelevant; rather, it highlights the impossibility of separating the definition of technical problems from the social framework to which it is associated [5]. As has been identified with other healthcare technologies, a user's experience with a specific device doesn't depend exclusively on its functionalities, but on the interplay of people, places, procedures, and assistive or healthcare programs [9]. As such, the user's attitude towards a robot is constructed within

Table 1: A summary of each trial (T*i*) in the in situ session, showing the type of failure that was caused, when and how the robot produced explanations, and the focus of the discussion between the interviewer and participant.

	Failure	Explanations			Discussion Focus
		Before Task	**During Task**	**After Task**	
T1	None	Narrative explanation		Task outcome	Initial impressions, explanation length
T2	None	Expectation briefing	Progress updates	Task outcome	Explanation timing, expectations vs. narratives
T3	Grasping	Task confirmation	Progress updates	Task outcome	Comparing explanations (failure vs. non-failure)
T4	Navigation	Task confirmation		Narrative explanation	Failure explanations and handling
T5	None	Task confirmation		Narrative explanation	User's mental model of the robot's decision-making

the context of the actual interaction, rather than stemming from a pre-existing and coherent disposition [20].

Indeed, various characteristics can be identified during the design process that may have contributed to shaping participants' preferences. Elements of the in situ session contributed to create conceptual gaps between the real experience and the target application. For example, some participants did not identify themselves as the target population for such a robot, viewing the robot as a tool better suited to users with reduced mobility. Others felt disconnected from the experience due to the use of predefined standard objects rather than objects belonging to the participants themselves.

P7: *I honestly think it is very useful for people with mobility issues. [...] For me, nowadays, it has little use.*

Technical limitations can impact participants' perceptions in various ways. Its slow movements, the difficulty of adjusting explanation volume to an appropriate level, its tendency to get stuck in tight spaces and the use of predefined objects and not personal ones contributed to shaping the users' perception of the robot as clumsy, requiring a more dynamic and less receptive attitude toward it.

P1: *This is very slow [...] I was about to grab the mug. Because I'm nervous and I don't like slow things.*

The focus group sessions corroborated these observations suggesting that preferences were closely linked to how participants established contact and interacted with the robot. It became evident in both groups that the preference for brief explanations was associated with the perception of the robot as clumsy and slow.

P1: [FG1] *It shouldn't have to repeat when I ask for a glass and it says, "I'm going to get a glass." It should just go straight to get the glass or the cloth or whatever. In that sense, it is very slow.*

Emphasising the contextual nature of technological artifacts, Feenberg [10] proposes the instrumentalisation thesis. When a robot is introduced into a specific setting, what it is and the relations it enables depend on its use in a given environment. In that environment, the conditions of possibility of their use are configured, which may be different from those originally intended and those that were contemplated in their design [10]. From this idea, it is especially relevant to introduce users' perspectives in relation to care automation processes, because the relationships with the robot are shaped by factors both of the robot's design and its environment of use. The functional possibilities of the robots and the values and organisational context where they are introduced configure users' views, opinions, practices, and experiences.

Overall, the sessions suggest that experiences are shaped by interactions in context; which supports other findings on the use of social robots [8, 12]. Users do not merely reveal pre-existing preferences but construct them through interaction, negotiation, and reflection. Further research is warranted in order to understand the relationship between a user's context and their explanation preferences.

6 Conclusion

This work presents preliminary insights on explanation delivery derived from in situ and focus group sessions. Our key preliminary observation is a tension between the immediate preference for unobtrusive, simple explanations and the need for detailed, actionable explanations during failure scenarios. Crucially, these preferences were contextually constructed and unstable throughout the study. Indeed, preferences evolved during the course of the in situ sessions and, in some cases, between the in situ sessions and focus groups. Factors such as the perceived task relevance, robot attributes such as appearance, speed and volume, and the setting of the interview all contributed to these preferences.

For future in situ design sessions, we can make the following recommendations. Firstly, it is important to ensure that the robot's tasks are relevant to participants. While the object retrieval task was identified as useful, a wider range of assistive tasks could better engage participants, resulting in more personal investment in explaining and resolving failures. Related to this, it is crucial that the design sessions emulate the target application as faithfully as possible. While we have already taken considerable steps in this direction, conducting in situ sessions with the target population, it is clear that improvements can be made. A promising direction in this regard may be to conduct studies over extended periods of time [14]. This requires the robot to act autonomously and adapt to users' preferences, ideally with a natural language interface for more flexible interaction and instruction [16]. Progressing in these areas may help to ensure that failure explanations are useful and contextually relevant for older adults at home.

Acknowledgments

This work was supported by Horizon Europe MSCA grant agreement No 101072488 (TRAIL); by project CHLOE-MAP (PID2023-152259OB-I00) funded by MCIU/AEI/10.13039/501100011033 and ERDF, UE; by project ROB-IN (PLEC2021-007859) which was funded by MCIU/AEI/10.13039/501100011033 and "European Union NextGenerationEU/PRTR"; by the EU under project ARISE (HORIZON-CL4 -2023-DIGITAL-EMERGING-01-101135959); and by the AGAUR through a predoctoral research fellowship under the FI-STEP programme (2025 STEP 00177).

References

[1] Anas Abou Allaban, Maozhen Wang, and Taşkın Padır. 2020. A systematic review of robotics research in support of in-home care for older adults. *Information* 11, 2 (2020), 75. doi:10.3390/info11020075

[2] Mir Farooq Ali, David Chukwudi Nchekwube, Oleg Genova, Alessandro Freddi, and Andrea Monteriù. 2023. An Assistive Robot in an Indoor Scenario: the Stretch Hello Robot as Environment Organizer. In *IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering*. IEEE, 676–681. doi:10.1109/MetroXRAINE58569.2023.10405748

[3] Antonio Andriella, Carme Torras, Carla Abdelnour, and Guillem Alenyà. 2022. Introducing CARESSER: A framework for in situ learning robot social assistance from expert knowledge and demonstrations. *User Modeling and User-Adapted Interaction* (2022). doi:10.1007/s11257-021-09316-5

[4] Jenay M Beer, Cory-Ann Smarr, Tiffany L Chen, Akanksha Prakash, Tracy L Mitzner, Charles C Kemp, and Wendy A Rogers. 2012. The domesticated robot: Design guidelines for assisting older adults to age in place. In *ACM/IEEE international conference on Human-Robot Interaction*. 335–342. doi:10.1145/2157689.2157806

[5] Wiebe E Bijker. 2010. How is technology made?—That is the question! *Cambridge journal of economics* 34, 1 (2010), 63–76. doi:10.1093/cje/bep068

[6] Turgay Caglar, Zahra Zahedi, and Sarath Sreedharan. 2025. Excuse My Explanations: Integrating Excuses and Model Reconciliation for Actionable Explanations. In *ACM/IEEE International Conference on Human-Robot Interaction*. 729–737. doi:10.1109/HRI61500.2025.10973938

[7] Harriet R Cameron, Simon Castle-Green, Muhammad Chughtai, Liz Dowthwaite, Ayse Kucukyilmaz, Horia A Maior, Victor Ngo, Eike Schneiders, and Bernd C Stahl. 2024. A Taxonomy of Domestic Robot Failure Outcomes: Understanding the impact of failure on trustworthiness of domestic robots. In *International Symposium on Trustworthy Autonomous Systems*. 1–14. doi:10.1145/3686038.3686050

[8] Devleena Das, Siddhartha Banerjee, and Sonia Chernova. 2021. Explainable AI for robot failures: Generating explanations that improve user assistance in fault recovery. In *ACM/IEEE international conference on human-robot interaction*. 351–360. doi:10.1145/3434073.3444657

[9] Sara Donetto, Amit Desai, Giulia Zoccatelli, Davina Allen, Sally Brearley, Anne Marie Rafferty, and Glenn Robert. 2021. Patient experience data as enacted: Sociomaterial perspectives and 'singular-multiples' in health care quality improvement research. *Sociology of health & illness* 43, 4 (2021), 1032–1050. doi:10.1111/1467-9566.13276

[10] Andrew Feenberg. 2010. *Between reason and experience: Essays in technology and modernity*. MIT Press. doi:10.7551/mitpress/8221.001.0001

[11] Ferran Gebellí, Raquel Ros, Séverin Lemaignan, and Anaís Garrell. 2024. Co-designing Explainable Robots: A Participatory Design Approach for HRI. In *IEEE International Conference on Robot and Human Interactive Communication*. IEEE, 1564–1570. doi:10.1109/RO-MAN60168.2024.10731212

[12] Parag Khanna, Elmira Yadollahi, Mårten Björkman, Iolanda Leite, and Christian Smith. 2023. Effects of explanation strategies to resolve failures in human-robot collaboration. In *IEEE International Conference on Robot and Human Interactive Communication*. IEEE, 1829–1836. doi:10.1109/RO-MAN57019.2023.10309394

[13] Gregory LeMasurier, Alvika Gautam, Zhao Han, Jacob W Crandall, and Holly A Yanco. 2024. Reactive or proactive? How robots should explain failures. In *ACM/IEEE International Conference on Human-Robot Interaction*. 413–422. doi:10.1145/3610977.3634963

[14] Anastasia K Ostrowski, Cynthia Breazeal, and Hae Won Park. 2021. Long-term co-design guidelines: Empowering older adults as co-designers of social robots. In *IEEE International Conference on Robot & Human Interactive Communication*. IEEE, 1165–1172. doi:10.1109/RO-MAN50785.2021.9515559

[15] Anastasia K Ostrowski, Cynthia Breazeal, and Hae Won Park. 2023. How do older adults engage as robot co-designers? Rapid prototyping supported by lived experiences with technology. In *Expanding the frontiers of design: A blessing or a curse?* Routledge. doi:10.1201/b22630-24

[16] Benedict Quartey, Eric Rosen, Stefanie Tellex, and George Konidaris. 2025. Verifiably following complex robot instructions with foundation models. In *International Conference on Robotics and Automation*. IEEE, 1–8. doi:10.1109/ICRA55743.2025.11127418

[17] Vinitha Ranganeni, Vy Nguyen, Henry Evans, Jane Evans, Julian Mehu, Samuel Olatunji, Wendy Rogers, Aaron Edsinger, Charles Kemp, and Maya Cakmak. 2024. Robots for Humanity: In-Home Deployment of Stretch RE2. In *Companion of the ACM/IEEE International Conference on Human-Robot Interaction*. 1299–1301. doi:10.1145/3610978.3641114

[18] Wendy A Rogers, Travis Kadylak, and Megan A Bayles. 2022. Maximizing the benefits of participatory design for human–robot interaction research with older adults. *Human Factors* 64, 3 (2022), 441–450. doi:10.1177/00187208211037465

[19] Cory-Ann Smarr, Tracy L Mitzner, Jenay M Beer, Akanksha Prakash, Tiffany L Chen, Charles C Kemp, and Wendy A Rogers. 2014. Domestic robots for older adults: Attitudes, preferences, and potential. *International journal of social robotics* 6 (2014), 229–247. doi:10.1007/s12369-013-0220-0

[20] Lucille Alice Suchman. 2007. *Human-machine reconfigurations: Plans and situated actions* (third ed.). Cambridge University Press. doi:10.1017/CBO9780511808418

[21] Lennart Wachowiak, Andrew Fenn, Haris Kamran, Andrew Coles, Oya Celiktutan, and Gerard Canal. 2024. When do people want an explanation from a robot?. In *ACM/IEEE International Conference on Human-Robot Interaction*. 752–761. doi:10.1145/3610977.3634990

Received 2025-12-08; accepted 2026-01-12