A Taxonomy of Domestic Robot Failure Outcomes: Understanding the impact of failure on trustworthiness of domestic robots

Harriet R. Cameron
harriet.cameron@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Simon Castle-Green
simon.castle-green@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Muhammad Chughtai
muhammad.chughtai@beko.co.uk
Beko Ltd
Leicester, UK

Liz Dowthwaite
liz.dowthwaite@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Ayse Kucukyilmaz
ayse.kucukyilmaz@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Horia A. Maior
horia.maior@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Victor Ngo
victor.ngo@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Eike Schneiders
eike.schneiders@nottingham.ac.uk
University of Nottingham
Nottingham, UK

Bernd C. Stahl
bernd.stahl@nottingham.ac.uk
University of Nottingham
Nottingham, UK

ABSTRACT

Domestic robots are fast becoming an integrated part of daily life. In anticipation of increased uptake of robotic assistants in the home, researchers and designers must investigate what makes domestic robotic interventions trustworthy or untrustworthy as a matter of urgency. This paper explores the concept of failure in domestic robotics, using the case of a dishwasher robot, and its impact on trustworthiness. It asks what constitutes trust, what constitutes failure, and what are the impacts failure may have on service providers, users, and the robot itself. We present the findings from four workshops with robotics experts and potential end users. We show that failure is simultaneously complex and predictable and re-evaluate existing taxonomies of failure, applying them to the domestic sphere, thereby highlighting social and corporate facets of failure that are not currently represented. We also provide a new taxonomy of failure outcomes to highlight how failures can breach trust, and what effects that breach may have.

CCS CONCEPTS

• **Applied computing** → **Consumer products**; • **Computer systems organization** → **Robotics**; Robotic autonomy; • **Human-centered computing** → *Interaction design.*

KEYWORDS

Human-Robot, Interaction, Trustworthiness, Failure, Domestic, Robots

TAS '24, September 16–18, 2024, Austin, TX, USA

ACM ISBN 979-8-4007-0989-0/24/09
https://doi.org/10.1145/3686038.3686050

ACM Reference Format:
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 *Second International Symposium on Trustworthy Autonomous Systems (TAS '24), September 16–18, 2024, Austin, TX, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3686038.3686050

1 INTRODUCTION

Domestic robots are becoming increasingly integrated into residential and domestic settings as companions and assistants. They provide companionship and complete chores ranging from vacuuming, to feeding pets, to mowing the lawn and more. As the general public are more regularly exposed to these technologies, a contextual understanding of the potential impact such artefacts can have on users is increasingly vital. More specifically, it is essential to investigate what makes domestic robots trustworthy or untrustworthy in order to understand how that may affect uptake and adoption. To this end, this paper investigates trust through the lens of failure and risk; how domestic robots fail; how that failure is perceived, understood, and dealt with; and what impact failure has on user willingness to continue using these technologies. These are important questions that need addressing before resources, infrastructure, and funding are fully committed to realising domestic robot integration. This work ultimately lays the groundwork for a radical re-examination of failure and trust that goes beyond binary understandings of 'failure' and 'success' in order to look at degrees of failure. We present the findings of four workshops conducted with robotics experts and members of the lay public regarding potential failures of domestic robots. Specifically, we use a hypothetical domestic robot designed to help with the loading and unloading of dishwashers to encourage participants to base their evaluations in a potential reality. Drawing on existing taxonomies of failure, we highlight where applications of robotics in domestic environments are currently underrepresented in

current research and offer our own contributions to established taxonomies to bridge this gap. As such, this paper addresses the following research question:
How does failure impact the trustworthiness of domestic robots?

2 BACKGROUND

Introducing robots to the domestic setting presents multiple, complex concerns that must be addressed in order to ensure trust in the technology. Here, we present an overview of work related to robotic environments and potential issues that may arise. We then address failure, providing a number of potential definitions before exploring how previous works have tackled failure. Here, we find that research around robot failure is typically focused on measuring changes to user perception, rather than understanding why and how perceptions have changed. We then introduce failure taxonomies as a means of providing an overview of robot failure. Finally, we provide a definition of trust, and showcase an alternative form of taxonomy that deviates from the standard.

2.1 The Domestic Setting

Since the first service robots entered the home in the late 1990s, robots have increasingly integrated into daily life. Further advances have been made in this time through the combination of domestic chore robots and Internet of Things (IoT) devices, for both inside [28, 35] and outside [34] maintenance of the private home. Recent works by, e.g., Verne [34] and Schneiders, Kanstrup, Kjeldskov and Skov [28] have investigated how robot owners customise their robots and the environments in which they must operate, in order to make the robot more efficient. These works have identified that to prevent the robot from conducting actions that could be perceived as 'failures', domestic robots at times need additional information provided through internal device sensors, e.g., edge-sensors or lidar; as well as external sensors such as smart-cameras, GPS, or movements sensors. Schneiders, Kanstrup, Kjeldskov and Skov [28] go further to discuss how robot owners adjust not just their environments, but also the digital ecology in which the robot operates in order to prevent failures or breakdowns. However, despite the increasing adoption of robots to streamline domestic duties, robotic automation brings with it several 'ironies' [1] resulting in the need for a human-in-the-loop. For example, while a lawnmower robot may negate the need for the user to mow their lawn, it also introduces new tasks such as maintenance of the robot and disentangling robotic failures. As such, ensuring the suitability of environment through the creation of robot friendly, or robot inclusive, spaces is an increasingly prevalent body of work. Currently, Elara et al. [10] argue that Human-Robot Interaction research is predominantly taking the human perspective, however, as a complementary notion to the 'How to design robots' paradigm, our research community needs to further investigate the 'How to design for robots'. The authors present a set of four design principles leading to better design of domestic environments for robot inclusivity, instead of better robot design. Specifically, Elara, Rojas and Chua [10] suggest that robot designers should consider i) observability, ii) accessibility, iii) activity, and iv) safety, ultimately aimed at improving the ways in which robots operate in shared spaces, thereby reducing the likelihood of breakdowns and failures.

2.2 Failure in Robotics

'Failure' in itself is a contestable phrase. What constitutes as failure in any given robotics-led scenario is highly contextual to the environment, the demands placed upon the robot, and the perception of users [8, 15]. Several definitions have been generated over time that broadly describe a failure as when a robot does not deliver the service it is expected to within anticipated timeframes or parameters, for instance, "the inability of the robot or the equipment used with the robot to function normally" [8]. We adopt the broadly accepted definition of "a degraded state of ability which causes the behaviour or service being performed by the system to deviate from the ideal, normal, or correct functionality" [6]. Further controversy arises from the distinguishing of failures from causes of failure such as errors, and faults. Steinbauer [31] for example define the differences as: "A failure is an event that occurs when the delivered service deviates from correct service. An error is that part of the system state that can cause a subsequent failure. A fault is the adjudged or hypothesized cause of an error" [31]. Honig and Oron-Gilad [15] disagree, defining the differences as errors being events that lead to system states, which ultimately lead to failures. Beyond the difficulties of defining failure, there is much existing literature that explores the reactions of users to robotic failure. Typically this research stems from exposing users to 'failures' predetermined by the research team, and then measuring user responses through quantitative metrics such as the Service Robot Acceptance Model [12]; the Usability, Social Acceptance, User Experience, and Societal Impact model (USUS) [34]; the Multidimensional Robot Attitude Scale (MRAS) [22]; or the Godspeed model [2]. Each of these questionnaires measure different elements of response to try and understand the impact that robotic behaviours have on varying elements of the user's attitudes. The Godspeed, for instance, is one of the most widely used surveys for measuring attitudes to robots before and after failure. It uses semantic differential scales to measure anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety; facets chosen from a literature review of previously conducted surveys [2]. The MRAS, on the other hand, uses a seven-point Likert scale to measure 12 different facets of user response to robots generated from the results of a series of questionnaires and discussions conducted with participants exposed to robots. These 12 facets are: familiarity, interest, negative attitude, self-efficacy, appearance, utility, cost, variety, control, social support, operation, and environmental fit [22]. From evaluation of these methods being applied, we can see broad differences in attitudes based on demographic information such as gender (of both the user and the perceived gender of the robot [9, 30], cultural background [21], prior experience [11], and personality [2]. However, we also see findings that show more universal truths, such as that failure reduces trust between user and robot, and can reduce the likelihood of the user re-engaging with the technology [7]. Further, trust is also reduced if the user is told that a robotic failure is their fault [7]. These findings are useful in order to showcase the importance of understanding failures, however, detailed information regarding how different kinds of failures impact user is still limited. Research contributing to taxonomies of failure provide a useful starting point, as they allow us to understand the range of failures that may be present.

2.3 Taxonomies of Failure

Despite the difficulty of defining failure, many efforts have been made to classify robotic failure into taxonomies that showcase the different elements that may contribute to a robot failing, or the different kinds of failure that may occur. Often, these taxonomies begin with a distinction between human-made faults and physical faults [8]. These taxonomies will then go on to further subdivide failures, for example by distinguishing human-made faults by design and interaction faults [8, 31]. Physical faults may also be subdivided by severity level such as catastrophic and benign [8, 18]. Other taxonomies may consider social elements such as social norm violations [13, 32], human errors [23, 25, 26] or even communication and processing [6]. One of the most broadly adopted failure taxonomies is the Robot Failure Human Information Processing Model generated by Honig and Oron-Gilad [15], which classifies failure into two primary types; technical and interaction. Technical failures in this model are caused by hardware or software errors, and interaction failures are "problems that arise from uncertainties in the interaction with the environment, other agents, and humans" [15]. Building on many of the taxonomies described above, they further break down failure events by functional severity, social severity, relevance, frequency, condition, and symptoms. Aside from the comprehensive consideration of existing taxonomies, Honig et al's taxonomy benefits from being one of very few identified that is specifically applied in a later paper to domestic robots [14].

2.4 Failure and Trust

Similar to the struggle of trying to find a universal definition for failure, trust also lacks a widely accepted definition. However, it is generally accepted that "trust is a psychological construct, the experience of which is the outcome of the interaction of people's values, attitudes, and moods and emotions" [16]. Further, trust is commonly viewed as an expression of confidence between two individuals, such that during an exchange of any kind they will not be harmed or put at risk as a result of the other individual's actions [36]. They do not clarify what is meant by 'harm'. In the context of Human-Robot Interaction (HRI), trust can be separated into two categories: Performance-based trust and relation-based trust. The concept of performance-based trust is based on the robot's capacity to perform tasks effectively. In contrast, the notion of relationship-based trust is founded upon the emotional bond and perceived intentions between a human and a robot [19]. Tolmeijer et al. [33] builds upon these understandings of trust in HRI, honing the definition to mean "a person's willingness to rely on a robot to carry out its duties" [33]. They go on to define a taxonomy of failure from a Human-Robot trust perspective, categorising HRI failures into four distinct types: Design, System, Expectation, and User failures, and their impact on trustworthiness. In addition to this trust-related taxonomy of failure, a total of nine mitigation strategies are offered to aid in the trust repair with the user after an instance of failure; five mitigation strategies for robotic failure: Fix, Interaction Design, Explanations, Apology, and Propose Alternative, and four mitigation strategies for user failure: Ask the Human for Justification, Show Emotion, Involve Authority Figure, and Training [33]. Tolmeijer, Weiss, Hanheide, Lindner, Powers, Dixon and Tielman [33] argue the justification for such mitigation strategies originates in the inevitable need for autonomous trust repair. Enriching the taxonomy of failures and discovering the range of failures possible may be extremely useful in contributing to the development of autonomous trust repair models.

3 METHODOLOGY

Four workshops were conducted between June and November 2023; two with robotics experts and two with members of the general public. Workshops offer a flexible, exploratory, and participatory approach to research [17, 27] that bring together participants with something in common – be that working in similar fields, or sharing a common interest in a topic [24]. They can be used to produce a number of different 'texts' to be analysed including transcripts, physical objects, and notes. As such, they provide an excellent platform to explore nebulous and hard to reach concepts such as 'trust' and 'failure', providing a safe setting for participants to explore and negotiate what these concepts mean to them. Initially, one workshop was planned with each group (one for public and one for experts) but following the data collection it became clear that further validating the findings by repeating the workshops would allow for a stronger contribution. Experts were expected to work in fields related to the development or deployment of robots, including both industry and academia. Workshop 1 ran for 7 experts, workshop 2 ran for 7 members of the public, workshop 3 ran for 2 further experts, and workshop 4 ran for 9 further members of the public. In total, 9 experts and 16 members of the public were involved. Workshops lasted for three hours and each followed the same basic structure shown in 1.

Workshops were approved by the University of Nottingham Computer Science ethics committee (CS-2022-R52) and conducted at the Cobot Maker Space [1], University of Nottingham. Participants were thanked for their time with lunch and a gift voucher. Workshops were captured on Dictaphone to .mp3 and to video .mp4 as a back-up. Files were manually transcribed and anonymised by the lead author. Written material generated by participants was also collected including notes from the ideation exercise and suggestions of failures. Quotes are attributed to participants for the remainder of this paper as A# for expert participants and B# for public participants. Demographic information was not collected from participants as it was not deemed relevant to the findings at this stage, although future research may take demographic details of participants into account. Transcripts, notes, and 'prototypes' created in activity three were subject to reflexive thematic analysis (RTA) [3–5] [3-5] conducted initially by the primary author and validated by the wider research team as per standard RTA practice. RTA is an approach to qualitative analysis developed by Braun and Clarke and requires systematic, consistent analysis of texts through an overt epistemological lens. We utilise an interpretivist lens [29] that explores "how humans make meaning of their worlds" (p.2) as "interpretive methods are particularly advantageous for surfacing situated knowledge, reading silences in narratives and the reasons for them, and identifying tacit knowledge that underpins cooperation, conflict, and other relationships" (p.7). Reflexive thematic analysis lends itself perfectly to this analysis it allows for themes to be iteratively drawn out of the data, whilst critically reflecting

[1] https://cobotmakerspace.org/

Segment	Purpose	Activities
Introduction (20 mins)	To introduce participants to the project, the researchers, and each other. Establish shared, initial meaning of concepts such as 'failure' and 'robot'.	Each person (including facilitators) introduces who they are, any relevant experience they have, and their interest in the workshop.
Activity One: What is a Robot (20 mins)	To explore the complexities of defining 'robot' and to create a shared definition to take forward.	Facilitated discussion. Facilitators encourage participants to directly speak to one another, but also offer prompts to follow up on ideas or to provoke further conversation, for example "do you know anyone who uses a robot at home" or "do you think that voice assistants could count as a robot". Facilitators write down definitions on a large flipchart, which is left visible for the remainder of the workshop.
Demonstration (5 mins)	To showcase an example of a robotic arm that may be used in a domestic setting in the future. A UR31 (see Figure 1) completes two pre-programmed sequences in which it demonstrates its full range of movement and picks up plastic cups from a counter to stack them on top of each other.	
Activity Two: Initial Impressions (25 mins)	To encourage reflection and discussion around the robot such as emotional response, potential uses of the robot, interest in using robot. Also to begin to discuss the role of a dishwasher robot in the home.	Facilitated discussion. Facilitators encourage participants to directly speak to one another or make notes that they can share. Facilitators will allow participants to share initial reactions, and then will encourage participants to list the actions involved in using a dishwasher. Facilitators will write down each step as it is mentioned on a large flipchart, which will remain visible for the rest of the workshop. Facilitators then encourage participants to reflect on having a robot to complete the actions listed, again only interjecting when needing to follow up a line of thought, or prompt further discussion e.g. "what would make you want to use a dishwasher robot?"
Activity Three: Ideation (15 mins)	To gather information about what an 'ideal' dishwasher robot may look like, do, or behave.	Participants are given drawing tools (pens, paper) and basic crafting materials (pipe cleaners, straws, glue, string, elastic bands etc.) and are encouraged to draw, make, or write down what their perfect dishwasher robot would look like. Participants may choose to work alone or 'team up' with other participants. Once finished, they are asked by a facilitator to talk through their design choices. There is a 15 minute break scheduled for the end of this activity, so participants may choose to end early or continue crafting into the break.
Activity Four: Failures and Successes (30 mins)	To identify what would constitute a 'failure' or a 'success' in the context of a dishwasher robot.	Facilitated discussion. Participants are encouraged to reflect on all of the materials created in previous activities to identify what kinds of behaviours and actions would constitute the robot successfully undertaking a task, or failing to complete a task. One facilitator is guiding the conversation, and another is noting each failure or success on individual sticky notes. The person facilitating will encourage participants to discuss with each other, but will also offer prompts to follow up on ideas or to provoke further conversation, for example by drawing on behaviours that may have been mentioned in another part of the workshop.
Activity Five: Risk Matrix (25 mins)	To understand what the perceived likelihood and impact of each failure or success would be.	Facilitated discussion. Using the sticky notes created in the previous exercise, a facilitator will read out the behaviour to the group. Participants will then debate between themselves as to how likely the behaviour is to happen, and how impactful the behaviour would be if it were to happen. The end result is a populated risk matrix made up of the sticky notes.
Wrap up (10 mins)	To close discussion and capture any final thoughts	Participants are encouraged to reflect on the workshop and share any other insights that they have not yet shared. Participants are then debriefed and receive their vouchers in thanks for their time commitment.

Table 1: A list of workshop activities, their duration, and purpose.

Figure 1: A Universal Robots UR3 robotic arm demonstrating picking up and moving dishes

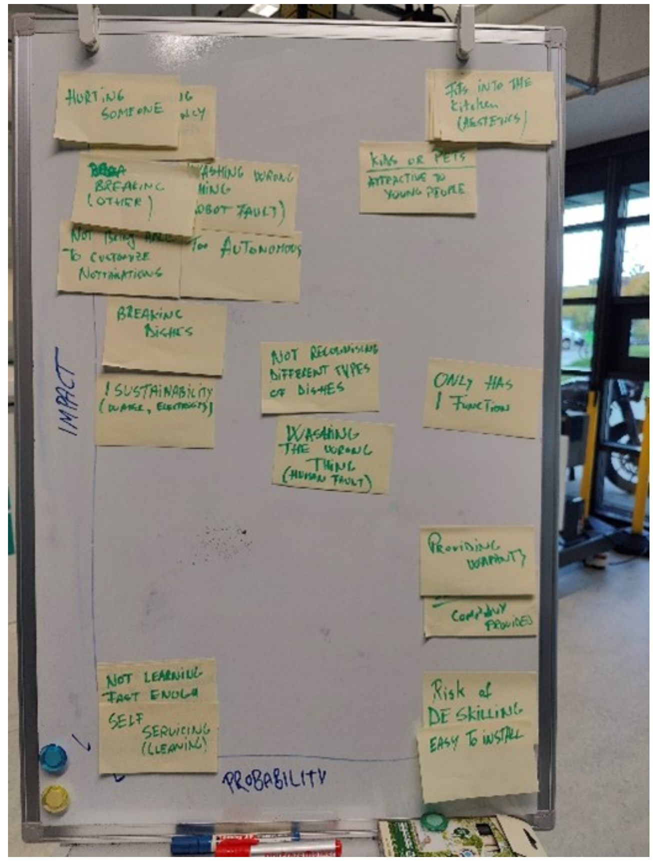

Figure 2: Risk Matrix from Second General Public Workshop

on what meanings are being represented; how we as researchers understand those meanings; and how our experiences, biases, and perspectives may shape the interpretations that we are making. As such, we are presented with a rich output of themes that reflect a complex tapestry of meaning-making from numerous perspectives, filled with nuance regarding the complex concepts of trust and failure addressed in the workshops.

Each activity was designed to feed into and provide references points for the next. Beginning by defining what a robot was, we collaboratively made a poster with a list of requirements as to what constituted, or did not constitute, a robot. In order to establish a shared understanding of an existing robots capabilities, participants were shown a series of pre-programmed demonstrations of the UR3 arm picking up and stacking cups. After the demonstration, participants were then encouraged to refer to the posters in order to discuss the robot demonstrated including their emotional responses, how it fit the definitions they provided, and whether they could see something similar within their homes. Following this discussion, participants were encouraged to ideate their own domestic robot. This could be based on the UR3, or something entirely different, but it should factor in the definitions and requirements discussed throughout the previous activities. Participants could take part in this activity either by writing their idea down, drawing it, or creating a model from arts and craft supplies provided. Participants then presented their robots and justified their design decisions to one another, calling back to previous discussions or drawing on their own personal experiences. Finally, participants were encouraged to draw on the material generated in all the previous activities including any models made, drawings, posters, sticky notes, or discussions, in order to identify potential failures of the robot. These were written on sticky notes, one failure per note, and participants ranked them on a risk matrix, failure impact on the x axis, and failure probability on the y axis. An example of a completed risk matrix can be seen in Figure 2. Written material from all activities were collated into an Excel spreadsheet for analytic purposes (see Table 2 for more details). Graphs from the final activity were translated to an excel spreadsheet where relative position on the graph was converted to a numerical score with 1 being low probability or impact, and 5 being high probability or impact. Risk scores were then generated by multiplying the probability factor and impact factor. Similar failure types across differing workshops were then consolidated for the purpose of generating average scores; for example, 'breaking dishes', 'damaging dishes', and 'breaking items' were consolidated as 'breaking dishes'. Where language used on written material was unclear, transcripts were cross-referenced to ensure accurate consolidation.

4 FINDINGS

Participants identified 44 potential failures in total (listed in Table 2). Risk scores for each failure can be seen in Table 2 with scores shared from each workshop, as well as the mean overall score. Scores are calculated by multiplying the 'probability' score (X-axis) with the 'impact' score (Y-axis). For the table and the remainder of the paper, workshops with members of the public are labelled P1 and P2 respectively, and workshops with experts are labelled E1 and E2 respectively. Failures are listed from low to high, and scores are further colour-coded for accessibility from light green (1.0-3.9; low risk), through dark green (4.0-4.9), yellow (5.0-9.9), orange (10.0-13.9), light red (14-15.9), and dark red (16.0+; high risk). This section unpacks the findings of the risk matrix activity through the evaluation of benign (1.0-4.9), moderate (5.0-13.9), and catastrophic (14+). Quotations are also provided from the workshops to highlight the context around failures where beneficial. Consensus around the trustworthiness, or likelihood of it completing its assigned task without putting the user at risk, of a domestic robot was mixed. There were high levels of uncertainty around many facets that contributed to this unease, however, there was also a level of expectation that the robot would, by default, be safe, considerate, and a good addition to the home: *"Well you'd trust it in some way to be a good companion in your home environment in some way, so it doesn't upset your cat or, you know, break your favourite plate. You know, there are many ways of breaching trust" (A3).*

4.1 Benign Risks

Five benign risks were identified (F1-5). Four of these risks were regarding the robot's capability to run correctly and autonomously (F1, 2, 4, 5) and were considered to be causes of minor inconvenience or mild *"irritation"* (B4). For instance, several participants agreed that cleaning and maintaining the dishwasher is challenging, with one even going so far as to describe topping up salt as *"the worst thing in the world"* (B16). Participants expected and wanted a domestic robot to accommodate for minor quality of life improvements that reduced burden around tedious or time-consuming tasks. These risks were usually seen, therefore, to impact the end-user if the robot did not accommodate for minor quality of life improvements, although that impact was generally small enough to not affect trustworthiness or use of the robot. One of the benign risks identified was that service providers would not offer a warranty for the robot breaking down (F3). This risk was deemed benign as it was considered very unlikely to happen. Participants expected that a warranty would be provided as part of the service, and if it were not, this was seen as a risk that would affect the trustworthiness of the service provider, not the robot.

4.2 Moderate Risks

Moderate risks accounted for the vast majority of failures (F6-34) and broadly covered failures related to the efficiency and safety of the robot. End users typically identified these failures as higher risk than experts, with notable exceptions in emergency situations (F12, 15, 18, 25), which experts scored higher. Within moderate risks, there were three kinds of harm that were identified that could arise as a result of the failures – physical, emotional, and financial. Physical harm (F11, 25) included accidental, passive, or technical errors that led to the robot hurting the end user (physical harm to pets or children was ranked as catastrophic and will be covered in the next section). This ranged from the robot breaking down and catching fire (A8), to the robot *"kill[ing] someone"* (A2). Both were considered extremely unlikely, but with extremely high impact. Emotional harm (F6, 12, 23, 24, 26, 28, 29, 31) were seen as kinds of failures that would embarrass, frustrate, upset, scare, or even isolate the user. Again, these were all seen to be risks that would impact the user and reduce trust in the robot:

> *"I felt like people might feel a bit unfriendly about the design, like more than considering it as a dishwasher, they might think it's scary because this is going to the houses of normal people, not into a techy's house, so unless the person is really fascinated about technology from a user side, this [robot] is going to be a bit scary for them".* (B2)

Concerns for certain user groups were particularly prevalent around emotional harm, including *"older people"* (B6), young children, pets, people *"sensitive to noises"* (B1), and disabled people who were more vulnerable to having a negative experience with the robot. For example, B5 raised a concern that a robot assistant could be embarrassing for disabled people who may not want attention to be drawn to their additional needs. They reflected on this in Activity Three when designing their own robot:

> *"But I was really keen to make sure that it fitted with the look of the kitchen. So that if you've got someone who has a disability or whatever, it's not obvious". (B5)*

Financial harm (F8, 9, 16, 18, 32) was also a concern for many participants who questioned where the responsibility for financial loss would fall. Many participants felt that given the novelty of the technology, the service provider should be responsible for any financial harm arising from the use of the robot in the home, at least for the first 12 months:

> *"So I will rent this machine for a year, for example, to see if I will adapt and then any break that I don't know it causes or something the company will be responsible for that. But then if I pay the full amount of the price - if it's too expensive I can pay installments every month - after I pay the full price, the responsibility will be all mine". (A8)*

This kind of reassurance was seen to make a big difference to whether the end user trusted the service provider and the robot, although it was generally discussed as something that would be nice to have, rather than something that was essential. This was further evidenced by a number of participants saying that they would not trust a robot with their more expensive or precious belongings, at least in the short term:

> *A1: "This can also be higher risk if your dish is very expensive. I don't know, if you have crystal glass or something." A2: "It's a point, isn't it. Maybe it doesn't wash your fine Bone China."*

Other risks that weren't directly seen to cause harm but were still considered to be moderate were largely risks that were seen to cause irritation or inconvenience. These risks were usually a slightly higher level of risk as they may eventually lead to some of the

Failure	ID	P1	P2	E1	E2	Mean
Robot does not run the dishwasher autonomously when full	F1	1	-	-	-	1.0
Robot does not self-service/self-clean	F2	2	1	-	-	1.5
Manufacturer does not provide warranty	F3	2	-	-	-	2.0
Robot runs dishwasher 'incorrectly'	F4	3	-	1	-	2.0
Robot does not learn fast enough	F5	-	4	-	-	4.0
Human is deskilled	F6	-	5	-	-	5.0
Robot stacks dishes incorrectly or inefficiently	F7	-	-	5	-	5.0
Robot washes the wrong thing or damages property (user at fault)	F8	-	6	-	4	5.0
Robot takes too many resources to set up/configure	F9	9	1	-	-	5.0
Robot gets hacked	F10	-	-	-	5	5.0
Robot has an electrical failure	F11	-	-	-	5	5.0
Robot inappropriately communicates with user (content of communication)	F12	10	4	2	-	5.3
Dishwasher is only usable with the robot	F13	6	-	-	-	6.0
Robot does not fit aesthetically into the home of the user	F14	6	-	-	-	6.0
Robot has an insufficient emergency response	F15	4	5	10	-	6.3
Robot is energy inefficient	F16	8	3	9	-	6.7
User is unable to define how things should be cleaned and when	F17	8	-	4	8	6.7
Robot breaks dishes	F18	9	6	5	9	7.3
Robot cannot discriminate between dishes	F19	9	9	4	-	7.3
Manufacturer does not maintain machine	F20	-	10	6	-	8.0
Robot is unable to grip objects	F21	-	-	8	-	8.0
Manufacturer does not adequately train workforce	F22	-	-	8	-	8.0
User does not understand how the robot works	F23	-	-	8	-	8.0
Robot does not conform to user way of living	F24	15	4	-	6	8.3
Robot physically harms users	F25	10	5	15	5	8.8
Robot is not designed to work for different kinds of users	F26	9	-	-	-	9.0
Robot requires too much space	F27	9	-	-	-	9.0
User does not feel in control	F28	-	-	9	-	9.0
Robot communicates inefficiently or inappropriately (mode of communication)	F29	10	5	15	-	10.0
Dishes come out dirty and robot does not take action	F30	16	-	6	8	10.0
Robot is not easily reprogrammable	F31	10	-	-	-	10.0
Robot washes the wrong thing or damages property (robot at fault)	F32	-	8	16	10	11.3
Manufacturer does not gather data needed to improve product	F33	-	-	12	-	12.0
Robot does not accommodate user requests	F34	20	5	-	-	12.5
Robot does not meet user expectations	F35	12	-	16	-	14.0
Robot is not convenient enough for users	F36	15	15	-	-	15.0
Robot is unsafe around children/pets	F37	15	20	-	10	15.0
Robot is inefficient	F38	15	-	-	-	15.0
Robot doesn't distinguish clean from dirty dishes	F39	16	-	-	-	16.0
Robot breaches social trust	F40	-	-	20	-	20.0
Company does not take responsibility for mistakes	F41	-	25	-	20	22.5
Cost too high	F42	-	25	20	25	23.3
Company does not offer enough reassurance	F43	25	25	-	20	23.3
Robot causes emotional harm to users	F44	-	-	25	-	25.0

Table 2: Risk Matrix Scores of Identified Failures. Empty fields indicate that this particular failure type was not mentioned in the corresponding workshop.

harms described above. For instance, incorrectly stacking dishes (F7), not discriminating between dishes (F19), lack of manufacturer maintenance (F20), and inability to grip (F21) may all lead to breakages and subsequent financial harm. Again, these were all risks that were burdened by the end-user, and which affected the trustworthiness of both service provider and robot. Further, some concerns around safety and suitability also began to emerge in the moderate risk score. Safety was a particularly present discussion point in the expert workshops, where hacking (F10) and lack of training (F22) were both seen to be risks that were unlikely, but which would have huge impact on end users. Where these safety concerns were not raised in the public workshops, suitability of the robot was a far more prevalent discussion point, for example whether it would physically (F27) or aesthetically (F14) fit in the home.

4.3 Catastrophic Risk

10 catastrophic risk failures were identified (F35-44), which were primarily social or service provider-based failures. These risks showcase areas that are evidently both more challenging, and more important, to build trust in. There was a higher level of consensus around scoring for the catastrophic risks across expert and public groups, although the public tended to score risks slightly higher. Again, all three kinds of harm identified above were represented in the most severe risks. Most catastrophic risks were related to emotional harm. Despite many specific examples of emotional harm being identified and ranked as moderate risk, participants ranked the overarching concept of emotional harm to be severe. Further examples of emotional harm that emerged in this section (F35, 36, 38, 40, 43, 44) all revolved around uncertainty – uncertainty about what the robot will do, how the robot will behave, and who will be responsible for what elements all create high levels of assumed probability and impact, and low levels of trust. Financial harm was also represented in catastrophic risk, with uncertainty around the company taking responsibility for financial harm (F41), and the cost of the robot being 'too high' in comparison to the functionality (F42). Indeed, most participants suggested that they would be unwilling to pay much additional cost at all for a robot that only conducted one role. Instead, participants generally agreed that in order to pay more than standard dishwasher costs, they would want the robot to be able to perform multiple tasks including preparing and cooking food, putting away shopping, and cleaning up after itself.

> *"I'm just curious about the idea of using a robot arm to do these things cause if you're collecting dishes and putting it then it can do other things, it can hand you ingredients, it could feed you if you can't feed yourself. And once you have that capability, surely it has other applications as well, being well beyond loading the dishwasher. Perhaps you know more involved in the preparation of food, consumption and then the clean-up". (A3)*

Other catastrophic risks were regarding potential physical harm to children or pets (F37). This risk was considered the same level of impact as harming the end user but was deemed far more likely to occur. This was explained as being due to unpredictability of children and animals: *"like my cat loves to go inside the dishwasher. So all the time I have to say go away, go away, not here"* (A8); as well as children potentially thinking of the robot as a toy to be played with: *"So they would think it's a toy. My 6 year old. She's autistic as well, so she would just think it's a normal toy"* (B14). As such, there was some consensus that the robot, in order to be fully trustworthy, *"has to work perfectly all the time"* (A1).

5 DISCUSSION

As shown, potential failures identified by participants varied in terms of both probability and impact. Where overall risk scores were higher, this appeared to have the greatest impact on relation-based trust. Where overall risk scores were lower, this seemed to have the greatest impact on performance-based trust. Further, we see that more oblique failures scored higher on impact than easy to understand failures. This interesting dichotomy leads us back to Elara, Rojas and Chua's [10] principles of robotic design; observability, accessibility, activity, and safety. From our findings, it becomes clear that uncertainty was a major contributor to perceived risk of failure. The more uncertain participants were in how a failure might occur, why, and what it could mean for the participant, the more impactful the failure was scored. As such, it becomes important to embed transparency and clarity into the robot design. Elara, Rojas and Chua's [10] principles showcase four important elements of the robot that could improve trust if addressed with transparency and clarity. For instance, the clearer it is what activities the robot is expected to undertake, the more performance-based trust the user may have that it will not deviate. The clearer it is that *safety* is a core concern of the design, e.g. that safety procedures have been put in place to prevent certain failures from occurring, the higher the performance-based trust may be. The more focus on making observability and accessibility clear features, the higher relation-based trust should be. As part of improving transparency and clarity, it also became clear that the existing definitions of trust in HRI are vague and do not clearly explicate what is meant by 'harm' or 'risk'. We postulate that failures impact trust by causing *physical, emotional, or financial harm* to the user. We propose, based on our findings, an adapted definition of trust in HRI – that is, confidence in the ability for user and robot to co-exist without increased risk of physical, emotional, or financial harm. To unpack this further, this discussion section first presents an overview of how trust is assigned and understood between the user, the robot, and the service provider. We adapt Honig and Oron-Gilad's [15] existing taxonomy to include these additional sources of failure that impact trust. In doing so, we examine what impact failure may have on different stakeholders. Finally, we present a new taxonomy that lays out the result of failure causing a breach in trust, including potential outcomes of that breach (Figure 3).

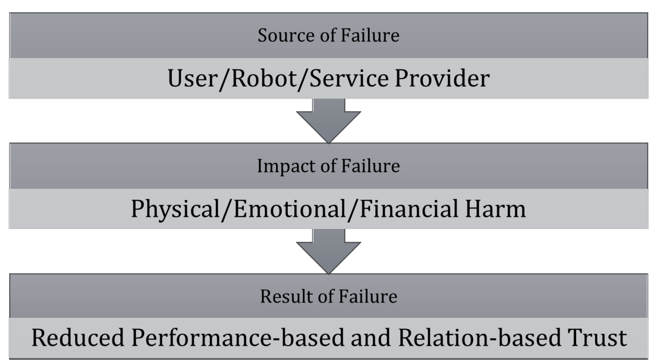

Figure 3: The Process of Failure

5.1 The Impact of Failure on Trustworthiness

As defined in [15], three sources of failure were identified; the user, the robot, and the service provider. Of the 44 identified failures, only one failure was placed at the feet of the user, which was allowing the robot to put something in the dishwasher that it shouldn't have (F8), for example, due to leaving a mobile phone unattended next to it. This is in some contradiction to the literature, whereby user failure was a common theme in the taxonomies, suggesting that users have higher levels of trust in their own performance, than in the robot or service providers'. In Honig and Oron-Gilad [15], failures allocated to the users are classified as slips, lapses, and mistakes. In the scenario presented to the participants, this may include examples such as failing to maintain the robot arm, failing to follow guidance set out by the service provider, or using the robot inappropriately. However, none of these possibilities were identified in any of the four workshops. This high level of trust the participants saw in end users and themselves may present some interesting challenges for designers and service providers in terms of ensuring correct and safe usage of domestic robots in the home. As shown by the risk scores, trust in the robot was the most variable. Baseline levels of trust were shown to be highly contextual, and to some degree, related to the trustworthiness of the service provider. The robots were assumed to be highly trustworthy to complete the tasks they were designed to do. However, as highlighted by [28] and [34], fitting into the environment and functioning consistently in a variable and changing domestic context was expected to present challenges, which, if failure occurred as a result, may have a lasting impact on trust. Further, it was also shown in the discussion that when a robot failed in any of the ways identified, this would have variable impact on the overall trustworthiness. Predictably, the higher the risk score, the higher the impact on trust was expected to be, with catastrophic failures shown to almost always affect relation-based trust, leading to the rejection of the robot. Moderate risks, however, mostly impacted performance-based trust, and were often anticipated to result in a change to the behaviour of the user, for example, by being more cautious in their interactions with the robot, or in altering how they would allow it to function. Benign risks were seen to have the greatest impact on the likelihood of the user to purchase a domestic robot in the first place, where if certain functionality was not offered, they would simply not buy one. For the service provider, trust levels were generally fairly low. Participants had some expectations about what the service provider would ensure was included with the robot, for example warranties and maintenance support, however, it was also expected that service providers would make it difficult to access these provisions. Half of the catastrophic failures were attributed to the service provider and would lead to either the user not purchasing the robot, or a complete rejection of the robot. As such, building trust in the service provider can be seen to be at least as important as building trust in the robot itself in order to ensure uptake.

5.2 Contributing to the Failure Taxonomy

In comparing our findings to existing taxonomies of robotic failure, we see much overlap in terms of types of failure identified within our workshops. Further, we also identify that the failures in existing taxonomies can be correlated to one of the three aspects of trust provided in our definition; physical, emotional, or financial. For the purpose of comparison here, we draw specifically on the taxonomy created by [14], as this taxonomy was created by drawing comprehensively on the work of existing taxonomies, and uniquely, is specific to the context of domestic robots. [14] break failure sources down into three aspects; service provider, user, and robot.

5.2.1 Service Provider Failures. The six failures attributed to service providers are shown in the flow chart captured in Figure 4 in white text boxes. Design is broken down into 10 further subsections. Within our workshops, we captured concerns from participants about all of the failures mentioned apart from delivery, including each of the subsections within 'design', validating and supporting the taxonomy provided. However, there were three further elements identified in the RTA that were not represented in the original taxonomy (presented in grey boxes), but which proved to be key aspects of failure within the workshops; business model, cost value, and training. Business model incorporates marketing (e.g., using misleading terms), additional requirements (e.g., requiring specific brands of periphery equipment, soft-locking other equipment when the robot is not in use/broken), and sustained support (e.g. the robot becoming obsolete if the company fails). Cost value was potentially a catastrophic failure, particularly in terms of the robot being too expensive to buy, the robot not being seen to provide enough value for its cost, or expensive maintenance and support. Training was also seen to be important, both for people related to the service provider (e.g. maintenance, installation, repair), and for the user (e.g. basic maintenance, safe usage).

5.2.2 User Failures. For user failures, [14] identify two kinds of failure shown in Figure 5, each with additional subtypes. These two failures are categorised as unintentional and intentional failures. Intentional failures were not seen within the workshops, by either our public or expert participants. Despite this, the category is well documented in the literature and is extremely important to consider in the design and application of robotic technologies. The fact that it did not emerge in our findings may even contribute to claims around the importance of this aspect, as if users are not aware of this potential failure, it may lead to adverse behaviour. Unintentional failures are subdivided into expectation, mistakes, lapses, and slips. Again, participants did not discuss many of the ways in which their actions may unintentionally lead to failures, highlighting an important consideration for designers. However, there was a reasonable amount of discussion across workshops specific to expectations. [14]'s taxonomy here provides commission (the robot does something the user does not expect) and omission (the robot does not act when the user expects) as potential failure points. These were both substantially recognised within the workshops, as participants tried to envisage how the robot would behave in the home. Here again, we also suggest three additional elements to be included in the taxonomy: adaptation, value, and anticipation of failure. Adaptation is where the user expects the robot to behave in a certain way that is adapted to the user's way of life. This differs from commission and omission in the sense that it is the robot neglecting to do something that the user incorrectly expects. Value and purpose, on the other hand, represents the users' 'failure' to correspond the value of the robot with the functionality it can provide. For instance, the user may anticipate that a robot has capability beyond its remit, and thus find themselves frustrated or disconnected from the robot when it does not live up to its anticipated potential. Alternatively, the user may refuse to engage with the robot, as they expect its value does not equate to its cost. Anticipation of failure was also a common theme highlighted and directly relates to low levels of trust in the technology to live up

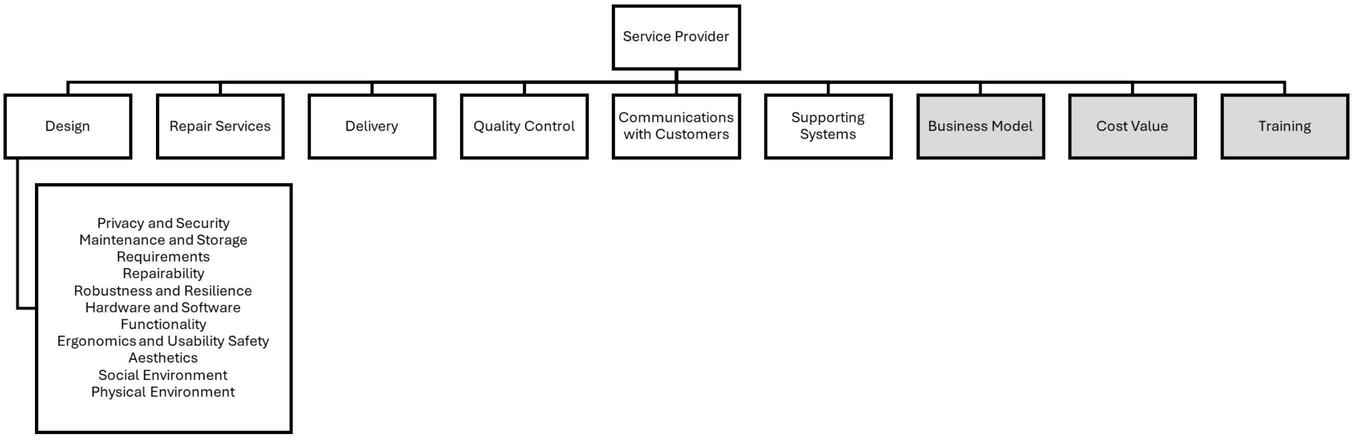

Figure 4: Failure Sources: Service Provider Failures

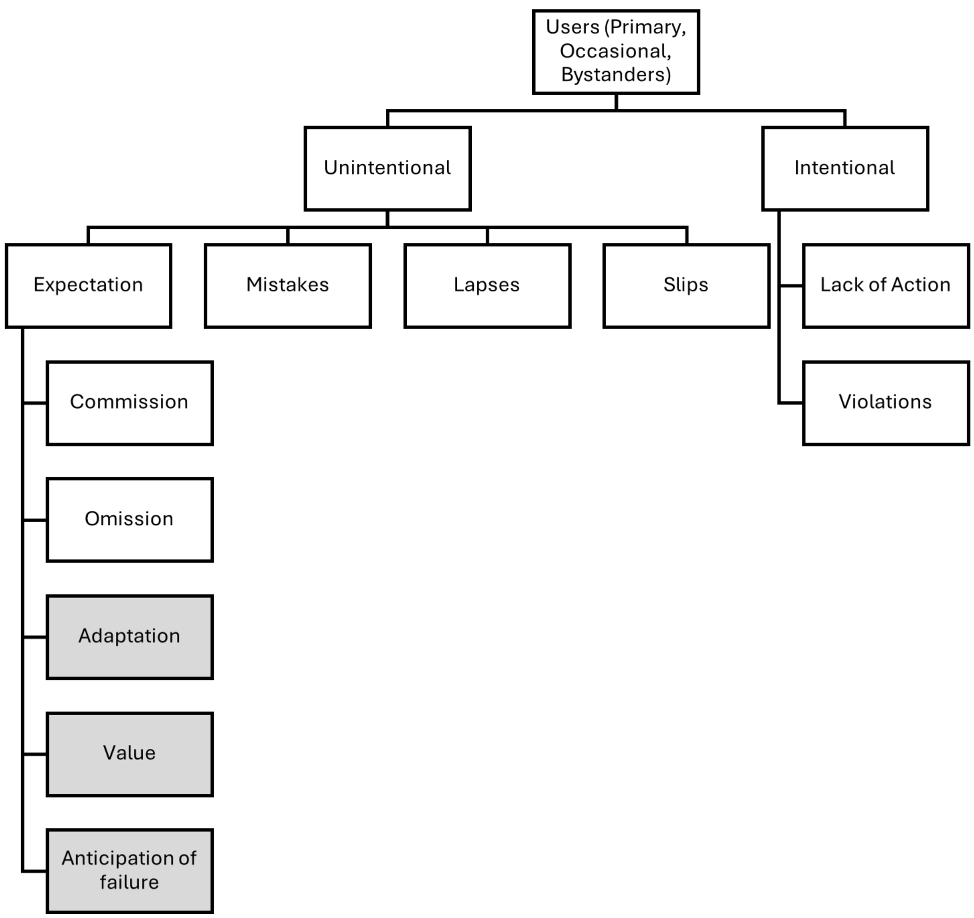

Figure 5: Failure Sources: User Failures

to expectation. Concisely, anticipation of failure is when the user does not allow the robot to complete certain actions as they do not trust it to succeed.

5.2.3 Robot Failures. Robot failures are broken down by [14] into technical and social failures, as seen in Figure 6. Some technical failures were discussed in the workshops, particularly by experts,

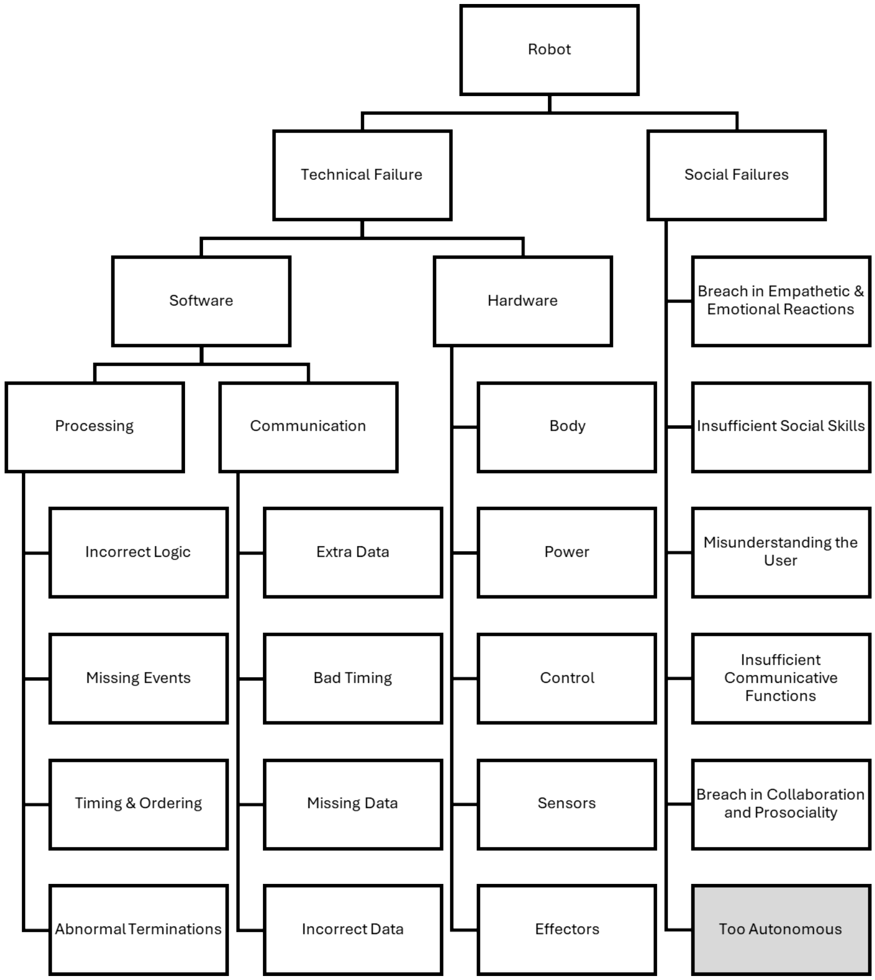

Figure 6: Failure Sources: Robot Failures

but the primary focus of this failure source fell into the social category. We observed discussions regarding all of the social failures described within the taxonomy across the full spectrum of risk. We have but one contribution to suggest for this branch, that of the robot being deemed 'too autonomous'. This occurs where the robot is seen to make decisions of its own volition without consulting the user, which directly reduces trust and increases 'creepiness'. In our example, this may include choosing when to put the dishwasher on without instruction, ordering supplies such as dishwasher tablets or salt without permission, or asking too many questions about instructions it is given. This was seen primarily to directly influence trustworthiness rather than risk. However, from the data and the literature, we also see that failures resulting in reduced trustworthiness eventually lead to cessation of use of the robot, and thus constitute as a failure under the definition provided.

5.3 A Taxonomy of Domestic Robot Failure Outcomes

Beyond our contributions to existing taxonomies of failure, our findings further revealed that failures were more naturally conceptualised and discussed in terms of the *outcome* of the failure, as opposed to the *causes* of the failure as captured in most existing taxonomies. Uncovering and tackling failures as they are understood by potential users is an important way of contributing to acceptance, trustworthiness, and uptake, particularly when considering the two different kinds of trust identified – performance-based and relation-based. Further, understanding how failures are conceptualised is an important part of the design process as it allows designers and programmers to better understand the underlying causes of failures that may not be related to hardware or software errors. As such, we also present a preliminary version of a *Taxonomy of Domestic Robot Failure Outcomes*:

The Taxonomy of Domestic Robot Failure Outcomes provides a step towards understanding the impact of different failures on the trustworthiness of domestic robots. Within each section, we

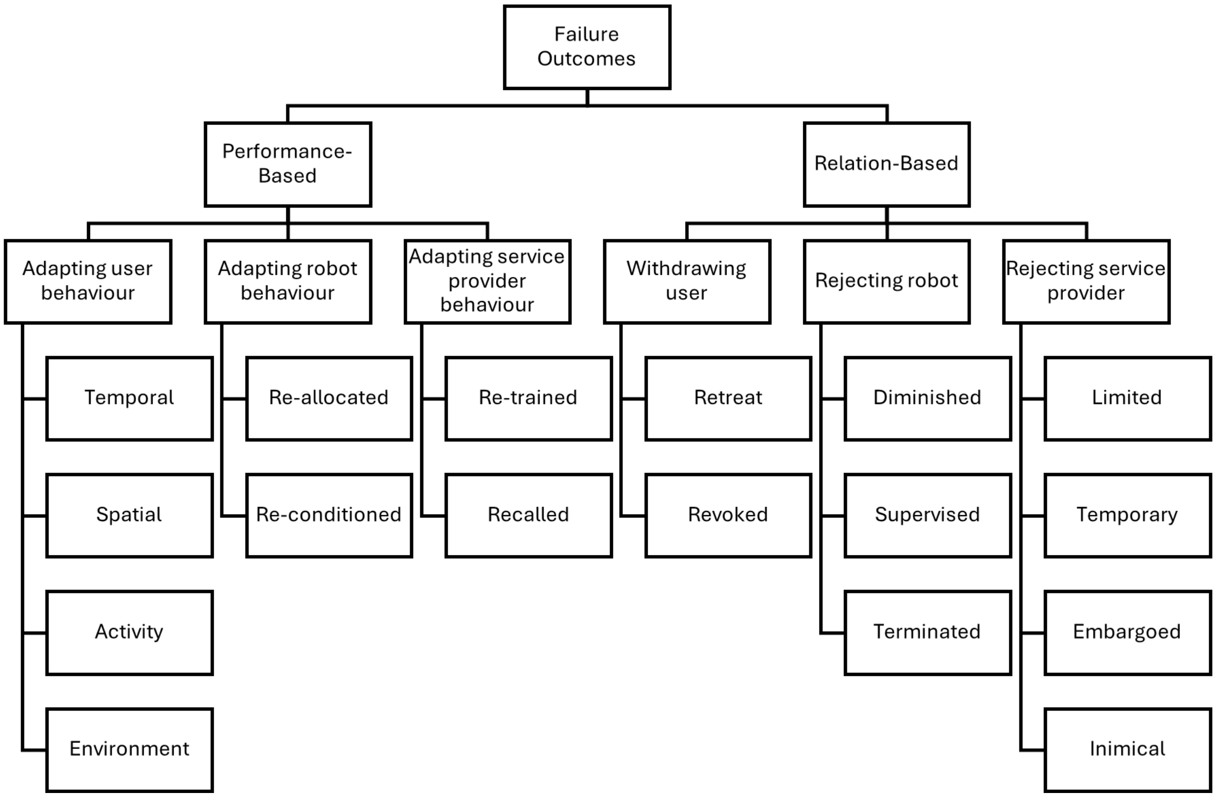

Figure 7: A Taxonomy of Domestic Robot Failure Outcomes

Failure Outcome	Subtype of Outcome	Definition
Adapting user behaviour	Temporal	Not acting when the user knows the robot is acting
	Spatial	Not acting within a spatial distance of the robot when it is acting
	Activity	Adjusting what activities are intended to be completed by the robot
	Environment	Adjusting the environment around the robot to be more suitable
Adapting robot behaviour	Re-allocated	Increasing or decreasing the workload of the robot.
	Conditioning	Changing the robot's programmed behaviours to better suit the needs and lifestyle of the user
Adapting service provider behaviour	Retraining	Service provider retrains workforce to deal with elements of the robot e.g. sales, marketing, maintenance
	Recall	Service provider recalls the service

Table 3: Definitions of Performance-based Failure Outcomes

present a potential outcome stemming from a failure. By utilising a lens of trust as defined by Law and Scheutz [20], we are able to divide impact of failures into behavioural (performance-based), and social (relation-based) impact. Further, we utilise the failure sources defined by [15] to divide failure outcomes by who the outcome affects. Subtypes were generated as part of the RTA and allow each outcome to be further quantified in a fashion similar to the outcome taxonomies shared above. Definitions for each outcome (as illustrated in Figure 7) of the taxonomy are elaborated in Table 3 and Table 4.

5.4 Future Research

In this paper, we present twofold contributions. First, we contribute additional insights into potential failures of domestic robots. Second,

Failure Outcome	Subtype of Outcome	Definition
Withdrawing User	Retreat	User prevents specific parties from engaging with the robot
	Revoked	User prevents any parties from engaging with the robot
Rejecting Robot	Diminished	Robot is used irregularly or intermittently
	Supervised	Robot is only used when it can be directly supervised
	Terminated	Robot is no longer in use
Rejecting Service Provider	Limited	User would purchase a different service from service provider again in the future
	Temporary	User would purchase the same service again at some point in the future
	Embargoed	User would not purchase from service provider again in the future
	Inimical	User vocally rallies against the service provider, for example through poor reviews or negative posts on social media

Table 4: Definitions of Relation-based Failure Outcomes

we begin the process of understanding the impact different kinds of failure can have on the user, the robot, and the service provider. These preliminary findings should further be substantiated through future research that focuses on understanding failure impact on trust. [20], whose research on trust in HRI provides the basis for our taxonomy, concludes their survey of the Human-Robot trust literature with the statement, "there are virtually no studies that objectively measure a purely relation-based trust" [20]. As a result, there is a gap in understanding surrounding people's trust of social robots to perform purely social tasks without a clear performance goal. We concur that exploration into the social side of robotic failures and how it affects trustworthiness in users is a necessary step towards more general acceptance, trustworthiness, and uptake. [33] revealed that little to no research on mitigation strategies for user failures has been undertaken. In correlation, our findings suggest that end-users are given higher levels of trust than that of the robot or the service provider. This notion that failures are more often than not perceived by end-users to be the fault of the robot may be an interesting challenge to designers and service providers when designing mitigation strategies for user failure, whether they be unintentional or intentional.

6 CONCLUSIONS

In this paper, we have investigated failure in domestic robots. To do so, we took a trifold approach. First, we defined and evaluated different forms of domestic robot failures according to probability and impact. Second, we applied those failures to existing taxonomies in order to understand where those failures may originate from. Finally, we present the Taxonomy of Domestic Robot Failure Outcomes as a means to investigate the impact of failures. From this investigation, we suggest an adapted definition of 'trust' for domestic robots; confidence in the ability for user and robot to co-exist without increased risk of physical, emotional, or financial harm.

ACKNOWLEDGMENTS

We grateful acknowledge the UKRI Horizon Trusted Data-Driven Products (Grant ID: EP/T022493/1), the UKRI Trustworthy Autonomous Systems Hub (Grant ID: EP/V00784X/1). We also gratefully acknowledge Beko Ltd for their partnership in this project. Generated data is unavailable due to ethics processes.

REFERENCES

[1] L. Bainbridge. 1983. *Ironies of Automation.* Pergamon, 129–135. https://doi.org/10.1016/B978-0-08-029348-6.50026-9
[2] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics* 1, 1 (2009), 71–81. https://doi.org/10.1007/s12369-008-0001-3
[3] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (2006), 77–101. https://doi.org/10.1191/1478088706qp063oa
[4] V. Braun and V. Clarke. 2013. *Successful Qualitative Research: A Practical Guide for Beginners.* SAGE Publications, London. https://books.google.co.uk/books?id=EV_Q06CUsXsC
[5] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676x.2019.1628806
[6] Daniel J Brooks. 2017. *A human-centric approach to autonomous robot failures.* Thesis. Computer Science.
[7] D. J. Brooks, M. Begum, and H. A. Yanco. 2016. Analysis of reactions towards failures and recovery strategies for autonomous robots. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN).* 487–492. https://doi.org/10.1109/ROMAN.2016.7745162
[8] J. Carlson and R. R. Murphy. 2005. How UGVs physically fail in the field. *IEEE Transactions on Robotics* 21, 3 (2005), 423–437. https://doi.org/10.1109/TRO.2004.838027
[9] Julie Carpenter, Joan M. Davis, Norah Erwin-Stewart, Tiffany R. Lee, John D. Bransford, and Nancy Vye. 2009. Gender Representation and Humanoid Robots Designed for Domestic Use. *International Journal of Social Robotics* 1, 3 (2009), 261–265. https://doi.org/10.1007/s12369-009-0016-4
[10] M. R. Elara, N. Rojas, and A. Chua. 2014. Design principles for robot inclusive spaces: A case study with Roomba. In *2014 IEEE International Conference on Robotics and Automation (ICRA).* 5593–5599. https://doi.org/10.1109/ICRA.2014.6907681
[11] Leopoldina Fortunati. 2018. Robotization and the domestic sphere. *New Media & Society* 20, 8 (2018), 2673–2690. https://doi.org/10.1177/1461444817729366
[12] Laura Fuentes-Moraleda, Patricia Díaz-Pérez, Alicia Orea-Giner, Ana Muñoz-Mazón, and Teresa Villacé-Molinero. 2020. Interaction between hotel service robots and humans: A hotel-specific Service Robot Acceptance Model (sRAM). *Tourism Management Perspectives* 36 (2020), 100751. https://doi.org/10.1016/j.tmp.2020.100751

[13] Manuel Giuliani, Nicole Mirnig, Gerald Stollnberger, Susanne Stadler, Roland Buchner, and Manfred Tscheligi. 2015. Systematic analysis of video data from different human–robot interaction studies: a categorization of social signals during error situations. *Frontiers in Psychology* 6 (2015). https://doi.org/10.3389/fpsyg.2015.00931

[14] Shanee Honig, Alon Bartal, Yisrael Parmet, and Tal Oron-Gilad. 2022. Using Online Customer Reviews to Classify, Predict, and Learn About Domestic Robot Failures. *International Journal of Social Robotics* (2022). https://doi.org/10.1007/s12369-022-00929-3

[15] Shanee Honig and Tal Oron-Gilad. 2018. Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development. *Frontiers in Psychology* 9 (2018). https://doi.org/10.3389/fpsyg.2018.00861

[16] Gareth R. Jones and Jennifer M. George. 1998. The Experience and Evolution of Trust: Implications for Cooperation and Teamwork. *The Academy of Management Review* 23, 3 (1998), 531–546. https://doi.org/10.2307/259293

[17] Mike Kesby. 2005. Retheorizing Empowerment-through-Participation as a Performance in Space: Beyond Tyranny to Transformation. *Signs: Journal of Women in Culture and Society* 30, 4 (2005), 2037–2065. https://doi.org/10.1086/428422

[18] Jean-Claude Laprie. 1995. Dependable computing: Concepts, limits, challenges. In *Special issue of the 25th international symposium on fault-tolerant computing*. Citeseer, 42–54.

[19] Theresa Law, Meia Chita-Tegmark, and Matthias Scheutz. 2021. The Interplay Between Emotional Intelligence, Trust, and Gender in Human–Robot Interaction. *International Journal of Social Robotics* 13, 2 (2021), 297–309. https://doi.org/10.1007/s12369-020-00624-1

[20] Theresa Law and Matthias Scheutz. 2021. *Chapter 2 - Trust: Recent concepts and evaluations in human-robot interaction*. Academic Press, 27–57. https://doi.org/10.1016/B978-0-12-819472-0.00002-2

[21] Hee Rin Lee, JaYoung Sung, Selma Šabanović, and Joenghye Han. 2012. Cultural design of domestic robots: A study of user expectations in Korea and the United States. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 803–808.

[22] Takumi Ninomiya, Akihito Fujita, Daisuke Suzuki, and Hiroyuki Umemuro. 2015. Development of the Multi-dimensional Robot Attitude Scale: Constructs of People's Attitudes Towards Domestic Robots *(Social Robotics)*. Springer International Publishing, 482–491.

[23] Donald A Norman. 1981. Categorization of action slips. *Psychological review* 88, 1 (1981), 1.

[24] Rikke Ørngreen and Karin Levinsen. 2017. Workshops as a research methodology. *Electronic Journal of e-Learning* 15, 1 (2017), 70–81.

[25] Jens Rasmussen. 1982. Human errors. A taxonomy for describing human malfunction in industrial installations. *Journal of Occupational Accidents* 4, 2 (1982), 311–333. https://doi.org/10.1016/0376-6349(82)90041-4

[26] James Reason. 1990. *Human error*. Cambridge university press.

[27] Amanda Rogers. 2010. The need for practical workshops. *Area* 42, 1 (2010), 127–131. https://doi.org/10.1111/j.1475-4762.2009.00897.x

[28] Eike Schneiders, Anne Marie Kanstrup, Jesper Kjeldskov, and Mikael B. Skov. 2021. Domestic Robots and the Dream of Automation: Understanding Human Interaction and Intervention. https://doi.org/10.1145/3411764.3445629

[29] Peregrine Schwartz-Shea and Dvora Yanow. 2020. Interpretivism. https://doi.org/10.4135/9781526421036915455

[30] Katie Seaborn and Alexa Frank. 2022. What Pronouns for Pepper? A Critical Review of Gender/ing in Research. , Article 239 pages. https://doi.org/10.1145/3491102.3501996

[31] Gerald Steinbauer. 2013. A Survey about Faults of Robots Used in RoboCup *(RoboCup 2012: Robot Soccer World Cup XVI)*. Springer Berlin Heidelberg, 344–355.

[32] Alistair Sutcliffe and Gordon Rugg. 1998. A Taxonomy of Error Types for Failure Analysis and Risk Assessment. *International Journal of Human–Computer Interaction* 10, 4 (1998), 381–405. https://doi.org/10.1207/s15327590ijhc1004_5

[33] Suzanne Tolmeijer, Astrid Weiss, Marc Hanheide, Felix Lindner, Thomas M. Powers, Clare Dixon, and Myrthe L. Tielman. 2020. Taxonomy of Trust-Relevant Failures and Mitigation Strategies. , 3–12 pages. https://doi.org/10.1145/3319502.3374793

[34] Guri B. Verne. 2020. Adapting to a Robot: Adapting Gardening and the Garden to fit a Robot Lawn Mower. , 34–42 pages. https://doi.org/10.1145/3371382.3380738

[35] Meryl Ye, Eike Schneiders, Wen-Ying Lee, and Malte Jung. 2023. The Future of Home Appliances: A Study on the Robotic Toaster as a Domestic Social Robot. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 477–482. https://doi.org/10.1109/RO-MAN57019.2023.10309555

[36] Lynne G. Zucker. 1987. Institutional Theories of Organization. *Annual Review of Sociology* 13 (1987), 443–464. http://www.jstor.org/stable/2083256

Received 17 June 2024; accepted 29 July 2024