The House That Saves Me? Assessing the Role of Smart Home Automation in Warning Scenarios

MARKUS HENKEL, PEASEC, Technical University of Darmstadt, Germany
STEFFEN HAESLER, PEASEC, Technical University of Darmstadt, Germany
HIBA K. AL-NAJMI, PEASEC, Technical University of Darmstadt, Germany
FRANK HESSEL, SEEMOO, Technical University of Darmstadt, Germany
CHRISTIAN REUTER, PEASEC, Technical University of Darmstadt, Germany

As smart home technology becomes integral to modern living, researchers must consider safety aspects. While single-purpose devices alert users to specific dangers, integrating them within comprehensive smart home warning systems (SHWSs) offers new safety potentials by allowing actuators to respond to threats based on predefined protocols. Key questions include whether user preferences for automation levels in smart homes are affected by different warning scenarios, and how unwanted automation or false positives influence acceptance. To explore this, we conduct two studies: (1) A lab study in a smart home with various actuators, where participants ($N = 48$) encounter warnings across three automation levels. (2) A follow-up interview study ($N = 16$) further evaluating our prototype and unwanted automation situations. Results show that participants preferred higher automation during warnings and were more receptive to smart technology in dangerous situations, though customization remains essential to ensure acceptance. While higher automation levels reduced perceived interruption, some still preferred less intense warnings. Others preferred not receiving warnings of mild dangers, fully relying on automation. Finally, we find that specific safety protocols and handling of false positive alarms must be chosen carefully to avoid mistrust, users feeling a loss of control, and damage through unwanted executions.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: smart home, warnings, automation, user study, crisis informatics, level of automation

ACM Reference Format:
Markus Henkel, Steffen Haesler, Hiba K. Al-Najmi, Frank Hessel, and Christian Reuter. 2025. The House That Saves Me? Assessing the Role of Smart Home Automation in Warning Scenarios. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 1, Article 5 (March 2025), 32 pages. https://doi.org/10.1145/3712269

1 Introduction

Over the past decade, smart home technology has become integral to many domestic environments, well-illustrated by the steady increase in user base [77] and by projections that anticipate a household penetration rate of 33% by 2028 [78]. Built on *Internet of Things* (IoT) technology, smart homes use interconnected sensors and actuators to offer various advantages in everyday life: interactive smart speakers provide information and entertainment,

Authors' Contact Information: Markus Henkel, henkel@peasec.tu-darmstadt.de, PEASEC, Technical University of Darmstadt, Darmstadt, Germany; Steffen Haesler, haesler@peasec.tu-darmstadt.de, PEASEC, Technical University of Darmstadt, Darmstadt, Germany; Hiba K. Al-Najmi, hiba.al-najmi@stud.tu-darmstadt.de, PEASEC, Technical University of Darmstadt, Darmstadt, Germany; Frank Hessel, fhessel@seemoo.tu-darmstadt.de, SEEMOO, Technical University of Darmstadt, Darmstadt, Germany; Christian Reuter, reuter@peasec.tu-darmstadt.de, PEASEC, Technical University of Darmstadt, Darmstadt, Germany.

ACM 2474-9567/2025/3-ART5
https://doi.org/10.1145/3712269

smart lights adjust to music and movies, and smart thermostats optimize energy efficiency. Correspondingly, research has investigated various aspects of such devices, including, e.g., the everyday use of smart speakers [7, 76] or design for understandable [21] and shared lighting systems [86]. Even those without technical experts can create tailored home automation setups using user-friendly options like trigger-action programming [83]. While many devices enhance everyday life, others specifically address safety and security concerns [73]. For example, doorbell cameras notify residents of visitors and smoke detectors can alert emergency services or trigger other actuators inside the home. However, many existing devices focus on single hazards rather than holistic solutions and do not integrate official governmental warnings ("public warnings") next to warnings of local dangers.

This paper addresses the role of automated actions in *smart home warning system*s (SHWSs), i.e., comprehensive, *multi-hazard* warning systems that increase safety and security in domestic spaces. Adding smart homes to the mix of existing warning channels, e.g., sirens, TV, radio, or smartphone alerts, makes sense because not all warning channels always reach their target audience, even if the population is actively made aware of tests [37]. Additionally, warning apps still record low adoption rates [44, 45, 56, 70], TV and radio need to be turned on to be effective, and traditional sirens may be limited by the topography and development in rural areas [47]. Therefore, SHWSs can serve as an additional warning channel for anyone who owns smart lights or speakers [42]. More importantly, however, we see two key advantages of sophisticated SHWSs over existing warning channels: *1) more comprehensive warnings*, and *2) automated safety protocols*.

(1) Public warnings inform citizens about regional dangers like natural disasters or large-scale accidents, but are unsuited for local hazards. Urgent issues like nearby gas leaks may escalate before operators can issue warnings, while in-home threats remain unknown to authorities. By combining public warnings and local threat detection using home sensors [72, 73, 81], SHWSs can provide more comprehensive information.
(2) As smart homes can integrate various actuators, the role of SHWSs extends beyond warning residents and includes actively mitigating dangers, leading to the vision of *the house that saves its residents*. By executing automated safety protocols, such as closing windows if gas or smoke is detected nearby, SHWSs could protect residents even if they are unaware or unable to respond themselves.

Automated safety protocols in particular offer unique potential for SHWSs to protect residents from dangers. Of course, existing research has already dealt with many aspects of smart homes [10, 27, 95] and SHWSs [73]. Next to (long-term) research on users' experiences [19, 52, 53] and creative workshops to co-design new use cases for IoT-enabled devices [1, 8, 9], safety, security [20, 68], and privacy [28, 100] are also actively researched in the home context. Regarding home automation, an important area of recent work has focused on making the use of systems user-friendly, for example through intelligible predictions of smart homes' behavior [19]. Paradigms that enable practical home automation setups have also been designed in user-friendly ways, including the popular trigger-action programming (e.g., IFTTT) [83], tangible end-user programming interfaces [22], or in-situ programming [62]. As trigger-action programs rarely work correctly on the first try, this strand of research includes supporting users in debugging their recipes [23, 99] or even delegating the programming to *large language model*s (LLMs) [38]. While these advances allow home automation to become more widely spread, an important question is how high automation should be. User interaction is not always mandatory, as these systems can automatically execute specified protocols based on triggers that occur. However, high levels of automation may result in a lack of control, causing frustration and impacting adoption intention [6, 94]. At the same time, the goal of home automation systems is not the manual execution of tasks, and specifically in the context of dangers in the home, higher automation may be preferable to assure the safety of residents. However, would fully automatic SHWSs lead to maximum safety for residents? And how would potential errors impact such decisions? Finding the optimal level of automation for SHWSs is not trivial, and while some existing research suggests that semi-automatic systems are preferable [74, 75], research on users' automation preferences of SHWSs is sparse.

This poses the question of how much autonomy is appropriate for SHWSs, and whether the danger and urgency of warnings impact the level of automation. Additionally, false positives need to be considered, as the goal of protecting users may clash with the consequences of unwanted automation. After reviewing related work in Section 2, we investigate these questions by creating a prototypical SHWS and conducting two complementary user studies based on it (see Section 3). The prototype is capable of automatically closing doors and windows as well as controlling power consumption of devices and ventilation systems, and includes three different automation levels and four simulated warnings. Study I quantitatively records participants' perceptions after interacting with our prototype (see Section 4). Section 5 then presents a follow-up mixed-methods study, adding qualitative insights to participants' perception of it and investigating false positives and unwanted automation. We discuss our results in Section 6 and provide important traits of SHWSs. Finally, Section 7 summarizes our work. Our results contribute to the area of HCI research on SHWSs that can warn of and react to dangers. By evaluating our prototype and exploring which factors influence the preferred automation level, our paper informs the design of future SHWSs to increase residents' safety and security in their homes.

2 Related Work

In the following, we summarize related research on smart home safety and the needs of smart home users (see Section 2.1). We then look into research on appropriate automation levels (see Section 2.2) and describe the research gap we aim to bridge through our work (see Section 2.3).

2.1 Smart Home Safety and Users' Needs

While many smart home devices provide users with convenient functionality to enhance everyday life [64], the areas of safety, security, and privacy have become increasingly important as well [68]. Some authors refer to systems that can monitor and warn residents of dangers as smart home alert systems [80] or smart home monitoring systems [98]. However, as our work goes beyond monitoring and focuses on warnings in both critical and non-critical situations, we will use the term *smart home warning systems* (SHWSs) [42, 72] in the following. Recent research on SHWSs shows a variety of architectures [73], with both wired and wireless sensors being used [85]. Applications are manifold, including intrusion detection, fire detection, or gas detection. As Sarhan shows, these systems differ in the types of sensors and detection methods used [73]: Intruder detection can be based on motion, vibration, or contact, while fires can be detected by smoke sensors, temperature sensors, or flame sensors, or by applying computer vision to live camera footage. Using *machine learning* (ML) techniques, these hazards can be classified with high accuracy [11]. Other safety-critical applications in smart homes include health monitoring in the context of ambient assisted living [25, 39] or the use of photovoltaic systems to "survive" power outages [49]. A recent practical example of SHWSs is showcased by a Ukrainian citizen, who, in response to Russia's invasion of Ukraine in 2022, built his own DIY SHWS to keep himself safe from missiles during the war [26, 29]. While many scientists focus on specific use cases, such systems are typically not mutually exclusive. By integrating multiple hazard warnings into one comprehensive SHWS, residents' safety and security can be further improved [42, 72]. Additionally, SHWSs are not limited to sensors and detection methods for local hazards but can also include public warnings, such as flood, earthquake, or tornado alarms, which are often issued via Cell Broadcast or warning apps [45]. The number of warning app users has increased in light of recent crises, but as many citizens still do not use such apps [44, 56, 70], we believe that SHWSs can help fill this gap by combining local warnings and publicly issued alerts in one system, ideally equipped with actuators that react to dangers.

While a lot of research on SHWSs has looked into technical improvements, the users' perspective plays an equally important role. Home automation may be seen as a positive form of technology, but recent research has taken a more critical stance, showing through co-design workshops that non-experts may end up marginalized [92]. Other obstacles in the way of SHWSs include the users' knowledge of safety-relevant functionality, their privacy concerns, and the appropriateness of warnings, as well as the perceived usefulness, ease of use, and control over the system. Some users simply do not think about the ability of cameras to detect fires in their homes, and instead use them to monitor their pets or to spy on neighbors [41]. Others know about safety-relevant functionality but believe that it is too expensive or struggle activating it [5, 6]. Next to knowledge-based obstacles, the users' perception is also vital to consider when designing interactive systems. In this context, it is crucial to assure that warnings not only function correctly on a technical level but also meet the users' expectations. Some users may wish to customize the geographic targeting of warnings or may only wish to receive specific types of warnings (e.g., they may be uninterested in school or food product alerts) [44, 56]. Unnecessary or excessive warnings can lead to alert fatigue [14, 87], resulting in important warnings not being taken as seriously anymore. Haesler et al. investigate this phenomenon in the context of SHWSs by providing a taxonomy of warnings based on their criticality and conduct a field study to evaluate different warning channels, such as *Short Message Service* (SMS), light bulbs, and audio warnings [42]. Another key requirement is for residents to be able to understand their homes' behavior. Coppers et al. develop an algorithm that uses predictions and simulations to assess which rules will be executed and evaluate their approach with diverse participants, finding that it leads to better understanding and trust in the system's actions [19]. Past research has also shown that users need to trust that their privacy is respected, as excessive data collection leads to lower adoption of smart home technology [17, 51, 93]. Solutions include sensible default privacy settings and the development of clear indicators to show that data is being collected, similar to red lights next to cameras [100]. Other examples are haptic personal privacy assistants that can be used to control multiple smart home devices' sensors at once [24, 65] or tangible privacy dashboards to raise awareness of data collection in smart homes [90]. Solutions for developers can entail the use of privacy-aware design tools that consider privacy concerns from the very beginning of development [2].

2.2 Appropriate Levels of Automation

As discussed above, the large variety of actuators in smart homes means that SHWSs can react to anomalies in various ways. Warnings can be presented using different warning channels, informing both present and absent residents and even notifying emergency services. Residents can view cameras or otherwise interact with the system, both from inside their homes via built-in displays and from afar via smartphone apps. SHWSs can automatically activate water sprinklers to extinguish fires [3], adjust water valves [72], close doors and windows [50], turn on the lights when a security camera notices potential intruders, or control the power consumption of electronic devices. As these examples show, the level of automation can range from purely providing information to interaction-based systems to fully automatic homes. Similar to the vision of cities' infrastructures responding to emergencies [82], we believe that smart homes can protect residents on a smaller scale. However, as research in other safety-critical areas shows, finding the correct level of automation can be challenging. In fact, levels of automation have been researched in various domains for decades. Prior research has often used the term "level of automation", but as many taxonomies with differing definitions have been proposed in the past and research suggests creating one's own taxonomy based on the specific needs and requirements [84], we define our own three automation levels to clarify our studies' conditions (see Section 3 and thereafter).

Past research has looked into various aspects of automation, including the consequences on human performance [67]. For example, Endsley and Kaber have conducted studies on the *level of automation* (LOA) in dynamic control tasks, linking lower automation to better performance and intermediate automation to better situational awareness [31, 32, 55]. Others have investigated the effect of higher automation in production systems on the physical workload of workers, showing that semi-automated lines had several advantages over fully automated lines [4]. Next to the domains of manufacturing and production [35, 36, 54], levels of automation have also been investigated in aviation, where excessive automation can result in a change from active to passive information processing, leading to reduced situational awareness [16]. Another safety-critical area where automation plays an important role is the realm of autonomous driving. Rödel et al. found that user acceptance and experience are highest for levels of autonomy that have been deployed in modern cars, and decrease at higher levels [71].

Automation levels are also a concern in research on smart homes. Some studies have shown that excessive automation may not always be a positive thing. In fact, if the level of automation is too high, users may perceive a lack of control over the system, which significantly affects their attitudes and thus their adoption intention [6, 93, 94]. Low levels of automation, however, may not be ideal either. If users had to initiate every action of their smart homes, the perceived usefulness and ease of use may suffer, also leading to lower adoption [63, 66, 79]. For smart speakers, which are some of the most common smart home devices, finding the right amount of proactivity is a challenge that has been taken on in recent years [15]. Some solutions include ML to learn users' preferences [46], while other researchers use storyboard activities [97] or dialogue elicitation studies [88] to inquire into factors that influence smart speakers' desired proactivity. However, even for simple interactions, errors are still common [89]. Looking more generally into smart home automation, Zaidi et al. investigate the users' perspective on how current smart home systems deal with conflicts between routines, i.e., cases in which competing instructions are to be executed by the same device [96]. They conclude that user satisfaction of resolving these conflicts relies on context and that their personal values influence their expectations.

Looking more generally into smart home users' automation preferences, Schomakers et al. conducted a qualitative study and employed quantitative online questionnaires, finding that many users preferred semi-automated systems [74, 75]. This was likely the case because they retained a higher level of control, although factors like reliability and storage location also seem to play a role. However, their research was performed in the context of smart homes in general, while our focus lies specifically on SHWSs. Additionally, we wanted to investigate automation with a realistic prototype to gather more informed data than is possible with online experiments, because realism is an important factor in HCI research [60].

2.3 Research Gap

A considerable amount of research already exists on different aspects of smart homes, such as technical implementations, users' needs, or novel use cases of IoT devices. Additionally, research on levels of automation has been performed in various domains, including manufacturing, aviation, and autonomous driving. When it comes to smart homes, however, and specifically the notion of SHWSs, we note a lack of research on appropriate automation levels. An open question is whether there are differences in users' preferences of automation levels between critical and non-critical situations, and which factors influence the optimal level of automation. We aim to bridge this research gap by conducting two user studies in a realistic smart home, capable of warning residents and executing helpful actions during situations of varying criticality. We aim to answer the following research question:

RQ: Which role does automation play in smart home warning systems, and how do situations of varying danger and urgency influence the users' experiences with such systems?

3 Research Design

In order to answer our research question, **we created a prototypical SHWS inside a living lab and conducted two complementary user studies,** focusing on different aspects of automation in SHWSs (see Sections 4 and 5):

- Study I: We invited $N = 48$ participants into the living lab to interact with our prototype in varying situations, collecting quantitative data through questionnaires focused on stress, interruption, and control.
- Study II: To extend the results of Study I and cover missing aspects, we conducted a follow-up study with $N = 16$ participants in a mixed-methods design, gaining deeper insights into the perception of our prototype and focusing on false positives, unwanted automation, and factors influencing users' preferred automation levels.

4 Study I: Interaction With Our SHWS Prototype

This section describes the evaluation of our prototypical SHWS in a quantitative lab study. Section 4.1 begins by providing detailed information on our chosen methodology. Section 4.2 then presents the results of our questionnaires.

4.1 Methodology

This section details the methodology for Study I, most of which is also relevant for Study II. Section 4.1.1 begins by presenting six hypotheses about SHWSs, followed by our smart home setup in Section 4.1.2. Section 4.1.3 then provides information on the participants. Subsequently, Sections 4.1.4 and 4.1.5 elaborate on the chosen study design and variables, followed by participants' tasks and the study's procedure in Section 4.1.6. Ethical considerations are discussed in Section 4.1.7, after which Section 4.1.8 explains the analysis of the collected data. Results of Study I are presented in Section 4.2. Study II's methodology and results are presented jointly in Section 5.

4.1.1 Hypotheses.

- **H1 - High Danger ⇒ High Automation: Users prefer higher automation in smart home systems during more dangerous situations than in less dangerous, everyday situations.**
  Reasoning: While prior research has shown a preference for semi-automatic systems [74, 75], we expect that in high-risk situations, users will prioritize safety benefits provided by high automation, similar to how lane departure warning systems and lane keep assist are used in cars to prevent accidents.
- **H2 - High Danger ⇒ High Intent to Use: Users show a greater preference for utilizing smart home features during more dangerous situations than in less dangerous, everyday situations.**
  Reasoning: We expect users who are generally critical of smart home features to recognize the immediate safety benefits during dangerous situations and be more willing to utilize them during dangers.
- **H3 - High Danger ⇒ High Stress During Warning: The system's warnings during dangerous situations cause more stress than warnings of less dangerous, everyday situations.**
  Reasoning: As SHWSs should match the warning intensity to the criticality of the hazard [42], we expect that more dangerous situations in our study will lead to increased stress. This is a natural consequence of the system issuing appropriately intense warnings to ensure that users recognize the criticality.
- **H4 (bidirectional) - High Automation ⇒ Different Stress Level During Safety Protocol: The level of automation has an effect on the stress experienced while executing the safety protocol.**
  Reasoning: Higher automation could reduce stress by quickly handling urgent dangers, requiring less decision-making and avoiding panic situations. However, the sudden activation of multiple actuators might also overwhelm users or require undoing if disagreed with.
- **H5 - High Automation ⇒ Low Interruption: Higher levels of automation in smart homes result in less perceived interruption of everyday tasks than lower levels of automation.**
  Reasoning: We expect higher automation to cause fewer interruptions, as users will not need to manually safeguard their homes. Lower automation requires more manual input, causing more interruption.
- **H6 (bidirectional) - High Automation ⇒ Different Level of Control: The level of automation has an effect on the perceived control over the situation.**
  Reasoning: While prior research has shown that high automation can reduce perceived control [74, 75], highly automated SHWSs could help manage time-critical tasks, allowing users to focus on other tasks.

4.1.2 Smart Home and Setup. The study was conducted inside a tiny house on the campus of our university, which is used as a living lab for our interdisciplinary research team to investigate (socio-)technical solutions for crisis preparedness and mitigation (see Figure 1). It features various pieces of furniture and a kitchen, including a

(a) The building is covered with photovoltaic cells that provide electricity, allowing it to keep up the most important functionality during power outages.

(b) Interior: ceiling lights, LED strip, LED bulbs, sliding door, window, fan, smart speaker, flower pot, control panel. LED lights were used to issue warnings.

(c) While experiencing warnings, participants sorted colored cubes located on seven shelves, simulating a moderately challenging household activity.

Fig. 1. Outside (a) and inside (b + c) view of the smart home where our study took place.

wooden floor and ceiling. The interior comprises an open single room with several nooks and an open gallery. The installed photovoltaic system on the building's outer shell enables it to be self-sufficient during power outages and to keep up emergency operation of its systems. We refer to this as "island mode" and also include it in our study. Users can control the lights and various actuators, e.g., to adjust the building's sliding door or its windows.

In the context of our study, we created a SHWS prototype with a 3D printed and assembled modular open source smart speaker (see Figure 2a) that is based on the concept of Leinweber et al. [59]. It serves both as an input and output device, and can be used to react to warnings that are issued during our study. The speaker includes a button for user interaction as well as lights to communicate the system state. On a technical level, integration with the smart home system uses a Raspberry Pi configured with Mosquitto as a *Message Queuing Telemetry Transport* (MQTT) broker [48], alongside Node-RED[1] and Zigbee2MQTT[2] for communication and device control. The built-in components of the smart home, including the *light-emitting diode* (LED) ceiling lights, door and windows, were managed using the KNX to MQTT interface. For devices that were specifically installed for the purpose of our study, we connected a Zigbee dongle to the Raspberry Pi to allow for the integration and configuration of Zigbee devices. Each device was assigned a topic name and unique device name, which were used in Node-RED for precise device addressing and control.

The smart home already had a central control interface installed on a wall, which can be used to adjust the motorized sliding door (see Figure 2b), the motorized windows (see Figure 2c), and to manage the ceiling lights (see Figure 2e), all of which can also be controlled by our SHWS. In addition to the motorized process, the sliding door can also be opened manually by hand. In an effort to make the smart home more realistic and homely, we integrated various devices and items within the building, including specialized LED light bulbs, an LED light strip attached to the stairs' handrailing, a fan, and potted plants.

4.1.3 Participants. We recruited $N = 48$ participants (25 identifying as female, 23 as male, none as non-binary), aged between 18 and 32 ($M = 24.10, SD = 3.78$), through university-internal mailing lists and posters. We chose to recruit primarily students to reach more participants, as the mixed design of our study (see Section 4.1.4) meant that participants would be divided among the different factor levels of the between-subjects variable (i.e., automation level). 25 participants were studying in a Bachelor's program, 17 in a Master's program, and five

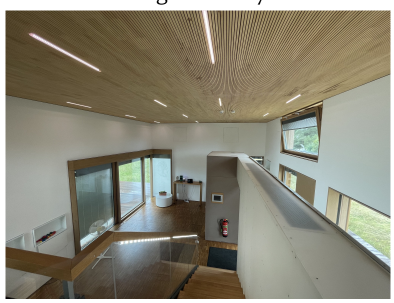
[1] https://nodered.org/
[2] https://www.zigbee2mqtt.io/


(a) Smart speaker and LED light bulb that were used to issue warnings.


(b) Sliding door that had to be opened or closed during the study.


(c) Windows that had to be opened or closed during the study.


(d) Fan that had to be turned on or off during the study.


(e) Ceiling lights that had to be turned off during the study.

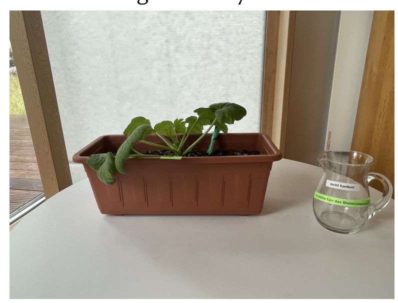

(f) Potted flowers that had to be watered during the study.

Fig. 2. The interior of the smart home, which not only has a kitchen and furniture, but is also equippable with various smart home interfaces that we used during our study.

were PhD students. To reduce sampling biases, we specifically recruited students from diverse majors, including thirteen participants studying architecture, nine studying computer science, and five studying psychology. Other majors included various engineering disciplines, political science, and sports science. Participants received 20 euros for approximately 45 minutes inside our smart home setup, ensuring compensation that exceeded the minimum wage for both their participation and travel time to our setup. Prior to the main study, we conducted two trials to ensure a smooth process.

4.1.4 Study Design and Independent Variables. Due to ethical concerns, we chose to conduct our study inside a living lab, allowing us to keep control over the setup. While realistic field studies would certainly provide more contextually valid data, we aimed to counteract the loss of realism through a distraction task and homely interior. We used a 3×4 mixed design in our study. **The first independent variable was the degree of automation of the SHWS, with three different factor levels (between-subjects).** Initial warnings were identical between all of them, as the SHWS always signaled a warning by activating the light bulbs and LED strip and by playing a sound through the smart speaker. However, further actions differed between the automation levels:

- **Automation Level 0 (AL0) - Low level of automation, manual execution.** If the user pressed the speaker's button, additional information about the detected threat and recommended actions were provided. The system did *not* automatically execute these actions; users had to perform them manually if they wished to implement the recommended steps.

- **Automation Level 1 (AL1) - Medium level of automation, timer-based execution.** If the user pressed the speaker's button, additional information about the detected threat was provided. Moreover, the system informed the user of the recommended actions and gave them a 30 second window to press the button again if they wished to abort the execution. If the button was not pressed within 30 seconds, the system automatically executed the actions of the predefined safety protocol.
- **Automation Level 2 (AL2) - High level of automation, immediate execution.** Without any user interaction required, the speaker immediately informed the user of the detected threat and which actions were recommended. The system instantly began executing the actions of the predefined safety protocol and informed users that they could press the speaker's button to cancel or undo them.

While interviews and online questionnaires have been used in prior research on smart home automation and found that users seem to prefer semi-automatic systems [74, 75], we wanted to investigate this topic through a study involving an actual prototype and focus specifically on smart home *warning* systems. We expect that higher automation may be particularly useful in dangerous situations, though the exact specifics of the system are not clear. Therefore, AL0 serves as an information-based system that warns residents, but requires interaction to provide additional details and does not execute any actions; participants have to execute the recommended actions on their own. For AL1 and AL2, we wanted the system to act automatically, but allow users a level of interaction to retain a feeling of control. AL1 requires some user input and then executes the safety protocol based on a timer, while AL2 does not require any input and executes the protocol immediately. The advantage of AL2 is quicker execution and not requiring user input, while AL1 may be seen as favorable when users disagree with the system's actions.

Our prototype's automatically executed actions include opening and closing the smart home's windows and doors, controlling the ventilation system, and disabling unnecessary electronic devices. For the warning of dry potting soil (see below), the system simulated watering the plants by announcing it via the speaker. Additionally, we attached a small pipe to the flower pot to make it look more realistic. Table 1 provides additional information on how each automation level affected our SHWS prototype in each of the simulated warnings, which are described next.

The second independent variable was the *warning scenario*, with four different situations (within-subjects) shown in Table 1. The specific warning scenarios were chosen because they represent typical dangers or problems that may occur in everyday life, including situations of varying criticality. The highest criticality is given by the detected gas leak, as it is both highly dangerous and urgent for residents to react to. Another warning was issued for intense heat, which can be dangerous to residents but does not pose the same urgency as a gas leak. As our third warning, a power outage typically does not pose immediate danger to residents, though it is still an urgent situation to be dealt with. Finally, the warning of dry potting soil neither poses a threat nor does it require an urgent reaction, resulting in a very low criticality.

We generally followed the recommendations by Haesler et al. [42], presenting critical warnings more intensely. However, we deviated slightly from their proposed taxonomy, and instead directly mapped the different attributes to modalities: *Danger* was mapped to light color, with higher dangers using red lights and lower dangers using white lights. *Urgency* was mapped to light permanence, with more urgent situations causing lights to flash and less urgent situations resulting in static lights. *Overall criticality* was mapped to sound, with critical warnings using sirens and less critical ones using notification sounds. We chose sensible recommended actions (which are also referred to as **safety protocol** in the following) for each warning scenario and included multiple actuators where possible.

As such, each participant was assigned one of the three automation levels and experienced four different warning scenarios.

Table 1. The four different situations that each participant was warned of during the study, including their danger, urgency, criticality, how they were presented, and the smart speaker's recommended actions.

Situation	Danger	Urgency	Criticality	Warning Modalities
Gas Leak	Very high	Very high	Very high	Siren sound, flashing red lights, speaker voice message "A gas leak has been detected in the area."
Recommended Actions: AL0: "Please close all windows and doors and turn off the fan." AL1: "In 30 seconds, I will close all windows and doors and turn off the fan. Push button to stop the execution." AL2: "I have to close the door, windows, and turn off the fan now. Push button to stop and undo the execution."				
Intense Heat	High	Medium	High	Siren sound, static red lights, speaker voice message: "Very high temperatures have been detected."
Recommended Actions: AL0: "Please open the door, all the windows, and turn on the fan." AL1: "In 30 seconds, I open the door, all the windows, and turn on the fan. Push button to stop the execution." AL2: "I will open the door, all the windows, and turn on the fan. Push button to stop and undo the execution."				
Power Outage	Low	High	Medium	Notification sound, flashing white lights, speaker voice message: "Strong fluctuations in the power grid were detected. House is in island mode."
Recommended Actions: AL0: "Try to save electricity and turn off all unnecessary lights and devices." AL1: "In 30 seconds, I will turn off all unnecessary lights and devices. Push button to stop the execution." AL2: "I will turn off all unnecessary lights and devices. Push button to stop and undo the execution."				
Dry Soil	None	Low	Very low	Notification sound, static white lights, speaker voice message: "Potting soil is very dry."
Recommended Actions: AL0: "Please water the plants." AL1: "In 30 seconds, I will water the plants. Push button to stop the execution." AL2: "I will water the plants."				

4.1.5 Dependent Variables and Control Variables. We collected responses to a mix of standardized questionnaires and custom Likert scale items that were specifically tailored to important aspects of our study. Unless otherwise noted, custom questions were based on 7-point Likert scales, with higher ratings indicating higher agreement. The full questionnaire is shown in Appendix B.1.

As dependent variables, we were interested in participants' automation preferences and intentions to use smart home features to find out whether there is a difference between everyday life and warning situations (H1 and H2). Additionally, we were interested in the stress that our prototype's warnings and actions would cause in participants, and whether the automation level had an influence on this (H3 and H4). We also assessed interruption and perceived control (H5 and H6). Additional questions were concerned with the perception of our prototype, including the *system usability scale* (SUS) [12] and the perceived suitability of the interfaces we used.

As control variables, we recorded the participants' perceived workload of a distraction task (see Section 4.1.6) through the raw NASA *Task Load Index* (TLX) [43] and their *affinity for technology interaction* (ATI) [34].

Additionally, we also recorded their prior knowledge of and experience with smart homes, which were recorded on 5-point Likert scales to facilitate comparison with related work [42].

4.1.6 Task and Procedure. The study was conducted over nine days, with each participant spending approximately 45 minutes inside the smart home. Upon arrival, participants were greeted, guided into the house, and given an overview of the study. After we collected consent forms and informed participants of their right to exit the study at any time, we emphasized that all warnings would be simulated and posed no real threat.

Participants were introduced to our SHWS prototype, experienced the speaker's voice output, and pressed its interactive button, which would later become relevant. To create more natural use contexts, we instructed participants to perform a distraction task, which consisted of sorting 200 colored cubes that were randomly distributed across seven trays (see Figure 1c). This task was designed to simulate attributes of common household activities - simple, repetitive, but requiring moderate concentration - to alleviate the loss of realism tied to any lab study. While participants sorted the cubes, warnings were issued in a randomized order using a Balanced Latin Square [30] to minimize unsystematic variance due to learning effects and carry-over effects. Participants had to resolve all situations on their own and were instructed to behave as they would in their own homes.

After each of the four warning scenarios, participants filled out a questionnaire about their interactive experiences, focusing on their perception of the system. Upon completing all warnings, a final questionnaire including ATI, general usability, and demographic questions was filled out, concluding their participation. Participants also filled out the NASA TLX questionnaire to assess whether sorting the cubes was sufficiently distracting.

4.1.7 Ethics. As our study encompassed warnings in fictitious danger scenarios and could therefore trigger traumatic responses or scare participants, we sought approval from our university's *institutional review board* (IRB) prior to the study. The approval was granted (Case EK 42/2023), and we made sure to clarify to participants that they had the right to terminate the study at any point, both during initial contact and at the beginning of the study. Additionally, participants were told that any warnings and dangers experienced during the study were not real. As explained in Section 4.1.3, monetary compensation was granted above the level of minimum wage.

4.1.8 Analysis. We used questionnaires to assess participants' perceptions of our prototype. Questions including more than one independent variable were analyzed using multilevel linear models and planned contrasts, as they have higher power than post-hoc tests [33], with details on our planned contrasts presented in Appendix A. Questions that only include one independent variable were analyzed using one-way independent ANOVA. As suggested by Field et al. [33], we only report significant main effects if the interaction effects are not significant. We use a significance level of $\alpha = .05$ for all statistical conclusions.

4.2 Results

The following section will present the results of Study I. Section 4.2.1 begins with the control variables, before Sections 4.2.2 and 4.2.3 focus on the hypotheses as well as the prototype's actions and interfaces, respectively.

4.2.1 Control Variables. We assessed the perceived workload of the distraction task using the NASA TLX and found that the task worked well, requiring moderate levels of attention without particularly high workload ($Mdn = 22.9$), comparable to everyday activities [40]. As all participants performed the same task regardless of their assigned automation level, there were no significant differences in their TLX scores, $F(2, 45) = 0.62, p = .542$.

Next, we controlled for participants' familiarity with smart homes on a scale from 1 to 5. While the term "smart home" was not new to most participants ($Mdn = 4$), prior experiences with smart home technology were less common ($Mdn = 2$), with 14 out of 48 participants never having interacted with any smart home technologies. We found no significant differences between the groups in terms of their familiarity with the term "smart home" $F(2, 45) = 0.85, p = .433$, or their prior experience with smart home technology $F(2, 45) = 0.33, p = .724$.

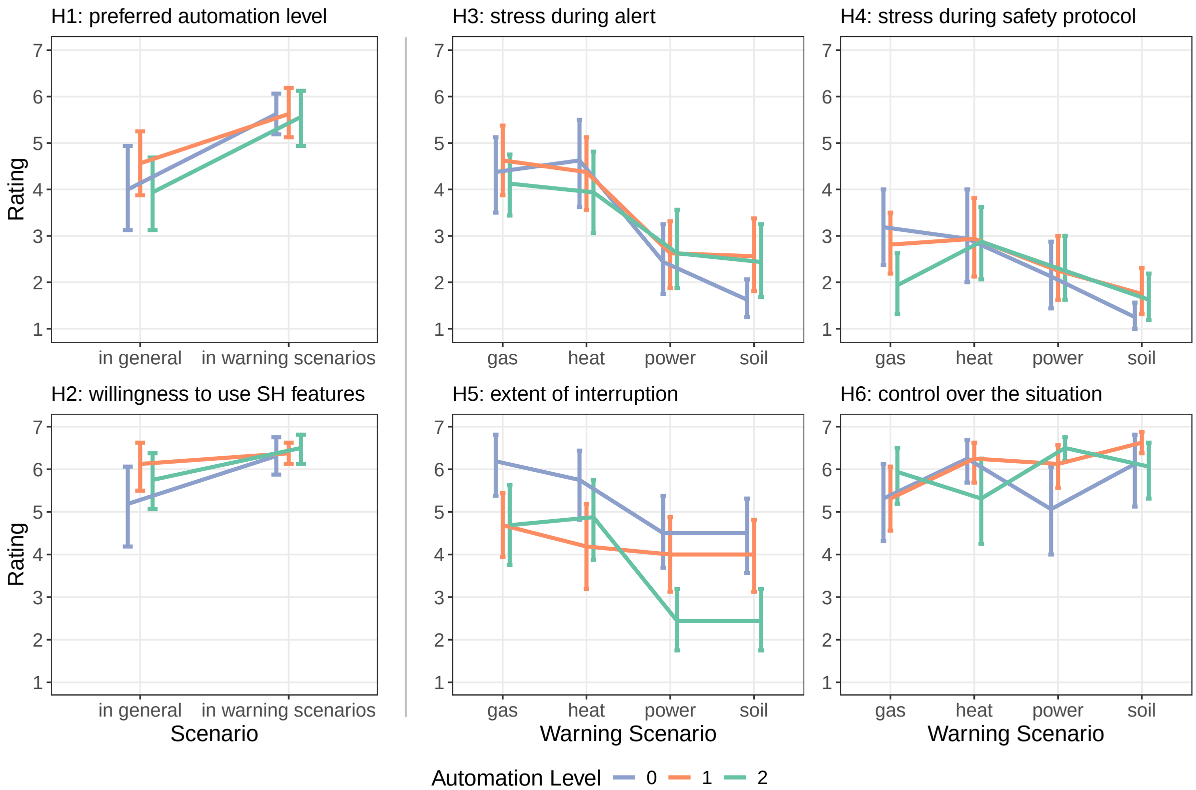


Fig. 3. Participants' ratings for the six Likert scale questions that were used to test our hypotheses. On the left: H1 and H2, which were answered after all scenarios occurred. Middle and right: H3-H6, which were answered after each of the four warning scenarios.

Finally, our participants showed a considerably higher affinity for technology interaction ($M = 4.5, SD = 0.8$, on a scale from 1 to 6) than the average ATI score of 3.61 found in a sample of German citizens [34]. No significant differences were found between the groups' ATI scores, $F(2, 45) = 2.76, p = .074$.

4.2.2 Hypotheses. Figure 3 shows participants' answers to the six Likert scale questions concerned with our hypotheses. Statistics are presented in Table 2 and allow us to summarize our findings as follows:

- **H1 - High Danger ⇒ High Automation:** Our results provide high statistical confidence to **accept** this hypothesis. Participants preferred significantly higher automation levels during dangerous situations than in everyday life.
- **H2 - High danger ⇒ High Intent to Use:** Our results provide high statistical confidence to **accept** this hypothesis. Participants were significantly more willing to use smart home features during dangerous situations than in everyday life.
- **H3 - High Danger ⇒ High Stress During Warning:** Our results provide high statistical confidence to **accept** this hypothesis. Participants experienced significantly more stress during warnings of dangerous situations than during warnings of less dangerous ones.

- **H4 - High Automation ⇒ Different Stress Level During Safety Protocol:** Our results do **not** provide statistical confidence to accept this hypothesis. We **reject** the hypothesis, as participants experienced similar, medium-to-low levels of stress regardless of the experienced automation level.
- **H5 - High Automation ⇒ Low Interruption:** Our results provide statistical confidence to **partly accept** this hypothesis. The significant interaction effect of automation level × warning scenario shows that the perceived interruption depends on the interaction of both independent variables. Breaking down this interaction, we find two significant contrasts, C6 and C9, which reveal two interesting effects:
  1) While participants always felt more interrupted using AL0 than using AL1, the difference between AL0 and AL1 was significantly more pronounced in dangerous situations than in less dangerous ones.
  2) While the perceived interruption was similar between AL2 and AL1 in dangerous situations, participants felt significantly less interrupted in less dangerous situations when using AL2 compared to AL1.
- **H6 - High Automation ⇒ Different Level of Control:** Our results provide statistical confidence to **partly accept** this hypothesis. The significant interaction effect of automation level × warning scenario shows that control over the situation depends on the interaction of both independent variables. Breaking down this interaction, we find one significant contrast, C10, which reveals an interesting effect:
  While participants using AL2 felt more in control over the gas leak than those using AL1, this effect was reversed for the intense heat scenario, where AL1 resulted in higher control than AL2.

4.2.3 System's Actions and Suitability of Interfaces. We also investigated other aspects of our prototype, including whether our prototype assisted participants well and whether the used interfaces were suitable.

We found a significant main effect of the warning scenario on participants' perception of the prototype's assistance, $\chi^2(3) = 33.12, p < .001$. Breaking down this effect, we find one significant contrast: C2, $b = 0.78, t(135) = 2.57, p = .011, r = .22$, tells us that participants felt like the assistance provided by the system was significantly better during the gas leak scenario than during the intense heat scenario.

Table 2. Summary of key statistics for each of the six hypotheses. "Smart Home" is shortened to "SH" in H2.

Hypothesis	**Main Effect** **Automation Level**	 **(Warning) Scenario**	**Interaction** **AL × WS**	**Significant Contrasts**
H1: preferred automation level	$\chi^2(2) = 0.86$ $p = .651$	$\chi^2(1) = 26.62$ $p < .001^*$	$\chi^2(2) = 1.24$ $p = .542$	-
H2: willingness to use SH features	$\chi^2(2) = 2.55$ $p = .28$	$\chi^2(1) = 10.75$ $p = .001^*$	$\chi^2(2) = 3.19$ $p = .203$	-
H3: stress during alert	$\chi^2(2) = 0.58$ $p = .748$	$\chi^2(3) = 81.53$ $p < .001^*$	$\chi^2(6) = 6.8$ $p = .339$	C3: $p < .001^*, r = .45$
H4: stress during safety protocol	$\chi^2(2) = 0.66$ $p = .72$	$\chi^2(3) = 25.48$ $p < .001^*$	$\chi^2(6) = 7.2$ $p = .302$	C3: $p = .012^*, r = .21$
H5: extent of interruption	$\chi^2(2) = 10.48$ $p = .005^*$	$\chi^2(3) = 50.27$ $p < .001^*$	$\chi^2(6) = 20.77$ $p = .002^*$	C6: $p = .019^*, r = .2$ C9: $p < .001^*, r = .35$
H6: control over the situation	$\chi^2(2) = 1.15$ $p = .564$	$\chi^2(3) = 8.5$ $p = .037^*$	$\chi^2(6) = 20.28$ $p = .003^*$	C10: $p = .01^*, r = .22$

* p < .05

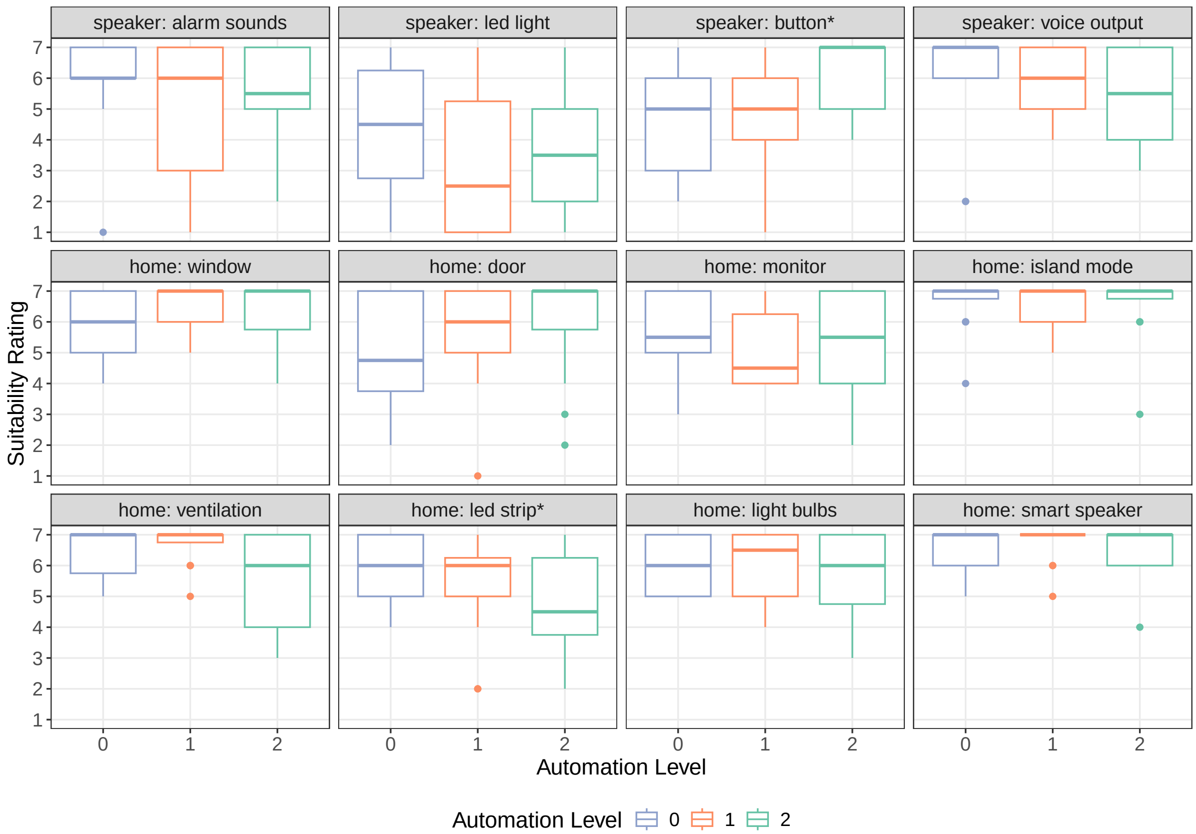


Fig. 4. Box plots for the perceived suitability of different interfaces. * denotes a significant effect of the automation level.

In an attempt to find out more about the individual interfaces that were used in our study, participants rated how suitable each of them was on a scale from 1 (not suitable at all) to 7 (very suitable). Figure 4 shows the results as box plots, split by the participants' experienced automation levels. Most interfaces were seen as suitable, including both common (smart speaker, ventilation) and uncommon smart home features (island mode for power outages, motorized door, and windows). The speaker's LED lights, however, were not rated as highly.

While the suitability of most interfaces did not vary between automation levels, there was an effect on the perception of the speaker's button, $F(2, 45) = 4.13, p = .023$. Using the same contrasts as in Appendix A, we find that C1 is not significant, $t(45) = -0.12$, $p = .909$, while C2 is significant, $t(45) = 2.43$, $p = .019$, $r = .34$. Therefore, participants that experienced AL2 considered the speaker's button significantly more suitable than those in AL1, but there were no substantial differences between AL0 and AL1.

We also found a significant effect of the automation level on participants' suitability rating of the LED strip (shown in Figure 1b) that was attached to the stairs' handrailing, $F(2, 45) = 3.27, p = .047$. While neither of our contrasts (C1 or C2) are significant, the respective box plot suggests that the difference lies in the fact that the LED strip was rated higher in AL0 ($M = 6.1, SD = 1.0$) than in AL2 ($M = 4.7, SD = 1.9$). Indeed, if we apply the more conservative Bonferroni correction, post hoc tests confirm this result, $p = .042$, $d = 0.92$.

5 Study II: False Positives and Unwanted Automation

While Study I provided valuable insights into the effects of automation in SHWSs, we conduct a follow-up study to cover missing aspects. In Study II, we collect qualitative data on the system interaction in several situations, including unwanted automation behavior, to more deeply analyze participants' perceptions of our prototype. We did not gather qualitative data in Study I to avoid it becoming too long. For the same reason, Study I focused on three distinct automation levels and four warning scenarios. Study II explores different situations where unwanted automation occurs due to false positives or varying individual preferences. We thereby address a critical aspect that we observed in Study I: the intense heat scenario was rated significantly worse and multiple participants pressed the smart speaker's button to cancel or undo the safety protocol. By explicitly focusing on unwanted automation, we aim to better understand the role of smart home automation in warning scenarios.

5.1 Methodology

To address these aspects, we conducted a **mixed-methods** follow-up study, gathering data on **the perception of our prototype, false positives, and unwanted automation**. We also investigated **more nuanced differences between automation levels** and **factors influencing them**. Additionally, we explored participants' **envisioned interaction** with SHWSs to avoid errors and mitigate their effects.

5.1.1 Participants. For Study II, we recruited a more diverse sample than in Study I, consisting of $N = 16$ participants with different backgrounds and ages (students, employees, but also retired, older adults, aged 19 to 67, $M = 32.07, SD = 15.66$). Eight identified as female, eight as male, and none as non-binary. We contacted them through an internal mailing list, a forum for user studies, personal networks, and on the street. Education levels ranged from Lower Secondary School degrees to Master's degrees and equivalent. ATI scores ($M = 4, SD = 1.07$) ranged from 2.44 to 6, resembling the national average more closely than Study I [34]. Four participants had already participated in Study I, while 12 were new to the study's context.

5.1.2 Procedure. Participation was possible both in person and remotely, and followed the same procedure regardless of modality to ensure equal conditions between participants. After collecting consent forms and clarifying the right to terminate the study at any point, we presented a brief overview of the procedure. We began by presenting photos of our smart home setup and explaining the context. We then showed participants video recordings of our prototype in each of the four warning scenarios used in Study I, randomized using a Balanced Latin Square. We chose an automation level between AL1 and AL2: when warnings occurred, the speaker provided information without the need to press a button, but the execution of the safety protocol was still based on a 30-second timer. After each scenario, participants shared their first impressions of our prototype's warnings and actions. Subsequently, we conducted semi-structured interviews pertaining to unwanted automation and errors (see Appendix C). Finally, participants filled out questionnaires (see Appendix B.2).

5.1.3 Analysis. To analyze our qualitative data, we followed a hybrid thematic analysis approach based on inductive and deductive coding using MAXQDA 24[3]. The process included three researchers (R1, R2, R3) and the following steps: 1) After R1 conducted and transcribed the interviews, R1 and R2 jointly created an initial codebook based on their domain knowledge. 2) R1 re-read all transcripts and refined the codebook, followed by a discussion with R2. 3) R1 and R3 used the codebook to independently code all transcripts, resulting in a substantial inter-rater reliability based on a Cohen's Kappa value of $\kappa = 0.7$ [18]. Disagreements were discussed and resolved. The final codebook used for analysis is shown in Appendix D. Questionnaire data was used in conjunction with the corresponding interview questions to better understand participants' perception of unwanted automation, envisioned interaction during errors, and interplay of factors that influence automation levels.

[3]https://www.maxqda.com/

5.2 Results

The following section will present the results of Study II. We begin with qualitative results of the perception of our prototype in Section 5.2.1, followed by Section 5.2.2 on errors and Section 5.2.3 about envisioned interactions, before Section 5.2.4 concludes with factors that influence the optimal level of automation in SHWSs.

5.2.1 Perception of Our Prototype. Results of our semi-structured interviews show that our prototype was generally well-received. Regarding perception of warnings, most participants commented on the different modalities, confirming that the speaker's sounds and the lights were perceived as intended. Sirens were clearly seen as more intense than the notification sound, as was the flashing of lights compared to static emission and the color red compared to white. However, many participants felt that warnings were generally too intense, especially for the heat and potting soil warnings. In case of dry potting soil, some participants stated that they would prefer no warning at all, and have the smart home water the plants automatically. Others only wanted to hear the speaker's notification sound or see a single light turn on near the flowers, omitting the voice message. The warning of heat was almost unanimously perceived as too intense, with some participants preferring lower intensity warnings, while others thought the warning was entirely unnecessary. As P5 put it:

> *"For the heat scenario, I wouldn't need either; I would figure it out myself. No notification, no action."* (P5)

Perceptions of our prototype's actions show a similar pattern, varying based on the warning scenario. For the nearby gas leak, participants spoke highly positively about the employed safety protocol. Few were unsure whether the fan should be turned off or remain on to spread gas that had already collected in the house evenly. Automatically closing windows and doors was seen as highly useful, especially as residents might not think about it during panic situations. P12 also highlighted the speed advantage of automation:

> *"I think the actions are good. It would be executed much faster, especially in a house or in a large apartment, if I didn't have to do it myself."* (P12)

Many participants were happy about automated watering of flowers, as they often forgot to do so on their own. For the power outage scenario, the safety protocol entailed a switch to the *island mode* to reduce energy consumption and keep up emergency supply through photovoltaics. Most participants appreciated turning off the fan, but turning off the lights was seen critically by some, as they worried about the risk of falling as a result of sudden darkness. Instead, a common suggestion was to *dim* the lights or turn off lights in unused rooms. The safety protocol for heat received by far the most negative responses from participants. Many were unsure whether windows should be opened at all, or whether they should remain closed. For example, P9 assumed the heat to be outside, and was concerned about it coming inside if windows and doors were opened by the SHWS.

> *"Depending on the situation, it makes less sense to open the door, as it might significantly heat up the room. An additional measure could be to have the shutters automatically go down when it's too warm outside."* (P9)

Even among those who understood the heat to be inside the home, concerns arose. For example, one participant viewed automatic opening of doors and windows as a security risk:

> *"I find the actions a bit questionable; I'm not sure if I'd want to open all my windows and doors for anyone. It's a security risk, after all. We can't assume that everyone has an ideal neighborhood or that there's no one within a 3-kilometer radius, and it's also a matter of privacy."* (P6)

While not all participants voiced concerns about the safety protocol for heat, the qualitative data from Study II confirms our suspicions from Study I and sheds more light on why the safety protocol for heat was rated significantly worse than for other scenarios.

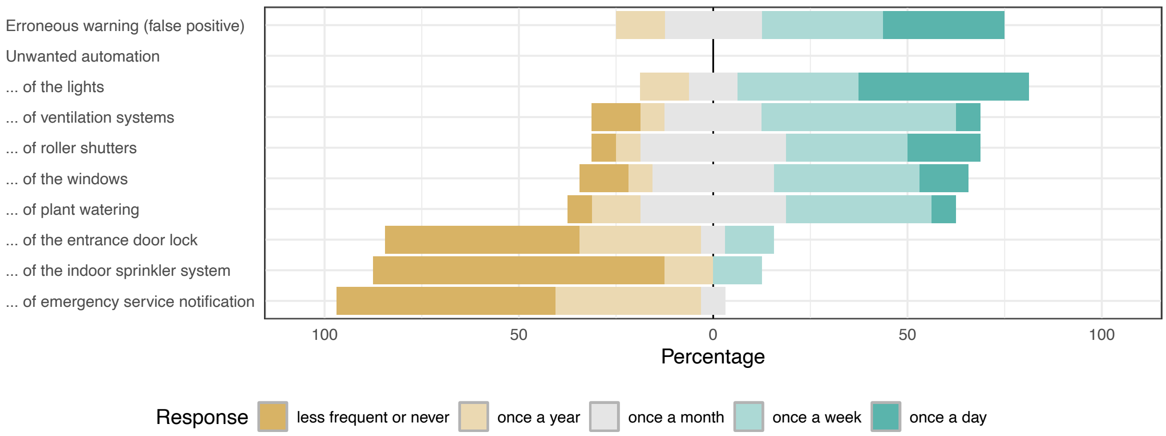


Fig. 5. Participants' responses to questions concerned with the maximum accepted frequency of errors.

5.2.2 False Positives and Unwanted Automation. Unwanted automation may play a crucial role in SHWSs and their acceptance. It can occur due to false positives (e.g., faulty sensors, public warnings), where the SHWS tries to protect residents from non-existent dangers. However, even in case of true positive warnings, i.e., warnings of existing threats, users may disagree with suggested safety protocols. Asked about their attitude towards errors, one person felt that unwanted automation would increase her feeling of safety, as it shows that the system is ready to protect her. However, most participants disagreed and reported unease thinking of such instances. Others feared that automation could cause harm. P4 remembered a movie she had watched and described the downsides of fully autonomous SHWSs:

> *"Recently, the movie "The House"*[4] *was on TV, which was exactly about a house like that, where a family lived. The house did things on its own, without them being able to control it, and I found that so oppressive and disturbing. I find that idea really awful. Seeing it visualized, how such a house would be, was super creepy."* (P4)

Next to fear and reduced safety, frustration and annoyance were common themes, for example if the user had to then undo the SHWS's action again. P3 described how even well-meant automation can cause resentment:

> *"You feel a bit patronized, and it could also become very annoying. For example, with the gas warning, if the window keeps opening and closing [...] I actually have personal experience with something like this. I once lived in a dorm that had automated shutters to save energy. Every morning at 6, the shutters in my room would move and wake me up, and there was no reason for it. If it happens regularly, it's simply frustrating and irritating."* (P3)

P13 wondered if SHWSs could identify false positives based on new data and undo prior actions. When he realized the problematic nature of re-opening doors at night after a gas leak had been identified as a false positive, his solution was that SHWSs only undo actions if false positives are identified within one hour of execution.

Some participants mentioned how false positive warnings could lead to reduced trust in the system. However, they understood that not all errors could be avoided and agreed that certain errors may be acceptable as long as they did not cause larger damage. Answers to our questionnaire reveal complementary trends (see Figure 5). Participants were more accepting of unwanted automation with low damaging potential (e.g., lights or ventilation) than of actions with high damaging potential (e.g., sprinklers or emergency service notification).

[4]2021 movie: a fully autonomous smart home develops a life of its own and threatens the owners, https://www.imdb.com/title/tt14917354/

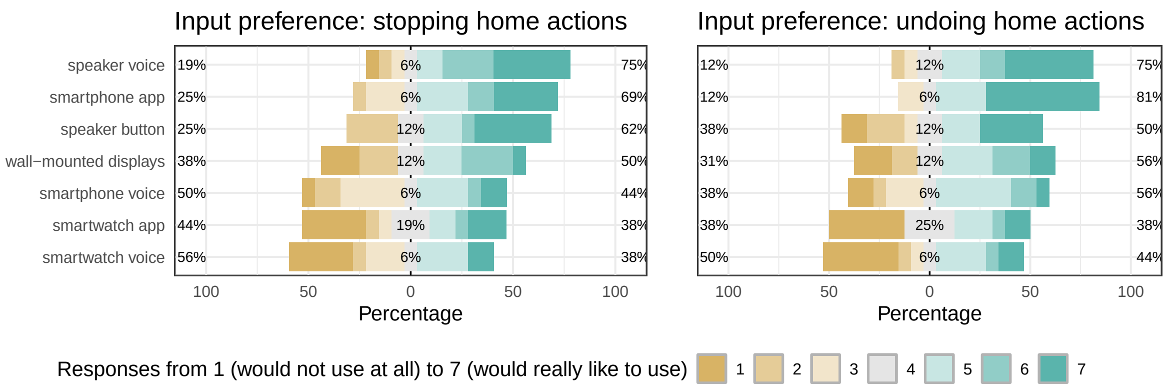


Fig. 6. Participants' responses to questions concerned with different input devices to stop or undo automated actions.

5.2.3 Envisioned Interaction with SHWSs Before, During, and After Errors. To avoid errors that cause damage, some participants wished for lower automation levels and for control over the system. This manifested in responses to one of our questions, which asked participants to choose between opt-in and timer-based opt-out systems, aiming to understand the considerations that affect the desired automation level. Interestingly, five participants favored opt-in automation, while eight favored opt-out systems. However, more than the remaining three participants struggled deciding on an option, as they saw advantages for each one. Reasons for selecting opt-in systems included avoidance of unwanted actions, forgetting to cancel actions when busy, and too high pressure of timer-based automation. Reasons for choosing timer-based opt-out systems included the reduced effort compared to always agreeing with actions, quicker execution, and the fact that opting-in is not always possible, as highlighted by P2:

> *"I would prefer [opt-out], because [opt-in] doesn't account for situations where there's a fire inside the house and I'm already unconscious on the floor. [...] There are certain dangers where I can no longer give active consent, so it would be bad if the house didn't respond on its own after a certain amount of time."* (P2)

Some participants also mentioned the importance of customization to make sure that their smart homes performed suitable actions. Such settings were typically envisioned as parts of smartphone apps, but also included dashboards usable on PCs or the ability to customize the SHWS via smart speakers. Desired customization options included fine-grained settings of automated actions for specific (types of) dangers, but also the ability to define how serious different hazards are in the user's opinion, as highlighted by P5:

> *"We've seen that the heat warning is not important to me. I'd like to be able to set up each individual case separately. It's hard to pinpoint a danger level, as the system might rate heat just as high as something more dangerous."* (P5)

When asked about the envisioned interaction with SHWSs before, during, or after unwanted automation, participants mentioned a mix of voice-based and touch-based ways to cancel or undo these actions. Most ($N = 12$) thought of talking to smart speakers, i.e., yelling "Stop!" (P1) or "Alexa, re-open the windows!" (P15). Smartphone apps were also commonly mentioned. Other devices included monitors mounted to walls and acting as control hubs, touchpads outside the house in case of being locked out, and buttons, e.g., on the smart speaker or walls. Figure 6 shows results of our questionnaire regarding interaction with SHWSs during or after unwanted automation, showing similar results.

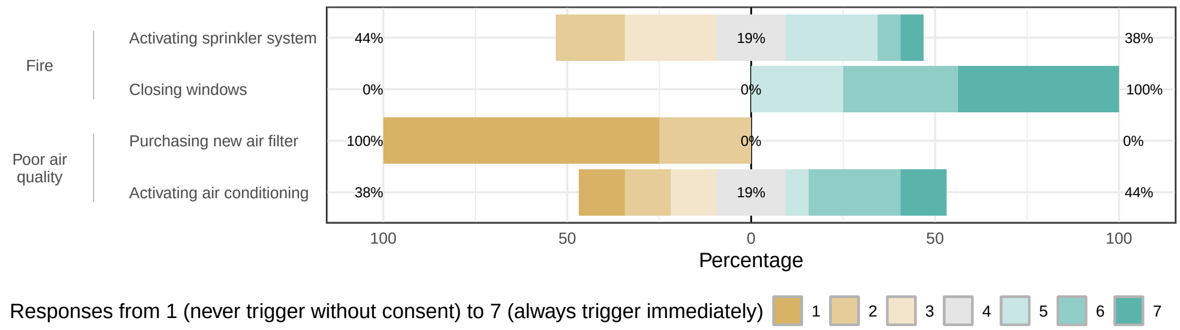


Fig. 7. Participants' responses to questions concerned with the optimal timer before SHWSs should perform actions in different suspected hazard scenarios. Hazards include one high criticality scenario (fire) and one low criticality scenario (poor air quality). Safety protocols entail two actions with highly negative consequences in case of false positive (sprinklers, purchasing air filter) and two actions with smaller negative consequences (closing windows, activating air conditioning).

5.2.4 Factors Influencing the Preferred Automation Level. We saw three key factors that may influence which level of automation is optimal for SHWSs: *danger*, *urgency*, and *negative impact of automation* during false positive warnings. To investigate how participants considered these aspects, our questionnaire assumed a *timer-based, opt-out* SHWS and presented two scenarios, each with two potential automated actions. Participants answered how long the SHWS should wait before automating each action, both on a Likert scale and via free text fields.

Figure 7 depicts participants' answers to these questions. Closing windows during suspected fires should happen almost immediately (Mdn = 7.5 seconds), while the purchase of air filters after detection of poor air quality should never happen without consent, or at least require high timers (Mdn = 24 hours). These results confirm our expectations, well-summarized by P1:

> *"In urgent or dangerous situations, I would like a high level of automation. In less urgent or less dangerous situations, I'd be fine with a lower or medium degree of automation to avoid damages caused by the system."* (P1)

However, these factors may be competing. Balancing the danger of fire and the damaging potential of sprinklers was more difficult for participants (Mdn = 2 minutes). Similarly, opinions on automated ventilation to improve air quality differed significantly (Mdn = 1 minute).

Next to danger, urgency, and negative impact of automation, our interviews also revealed a set of other factors influencing the optimal automation level of SHWSs. The full list of factors can be found as "Factors Influencing Automation Level" in the Appendix D. The following summarizes three different types of factors that we identified.

1) Personal and Contextual Factors include various aspects related to each individual person and situation. Some participants discussed age, and related to it, affinity for technology interaction and impairments, as influencing factors. For example, some participants highlighted how high automation could be particularly beneficial to individuals with physical impairments. Similarly, some imagined that elderly may benefit from higher automation, as having to stand up and manually "safeguard" the house could be delegated to the SHWS. Conversely, P7 thought that younger individuals may be more comfortable with higher automation due to growing up with technology. Living situation was another concern and entailed the size of the house or apartment as well other household members. For example, P14 struggled deciding between opt-in and opt-out systems because of this:

> *"I cannot really decide between these options. I was thinking, what if I'm not at home? Or what if it is not me consenting to the safety protocol, but my 7-year-old son? It's complicated."* (P14)

Finally, location emerged as an important factor. P6 preferred opt-out automation but thought that her SHWS should not be allowed to execute certain actions if she was not at home. However, P5 argued the other way round:

> *"If I'm not at home and the action can't be carried out because it's waiting for my active consent, that would be difficult. So I think it would be good to differentiate. If I'm at home, it should always ask for active consent, but if it's urgent, with a high level of danger, and I'm not home, then it should be able to execute without consent."* (P5)

2) System Design Factors play a crucial role in shaping automation levels, as they encompass features of the SHWS itself. For example, both the utility of automated actions and the effort that automation saves influence whether actions should be automated or not. Multiple participants also discussed that reversibility is an issue, highly related to the damaging potential of actions. Automating reversible actions, such as closing windows, was viewed as less problematic than irreversible ones, such as watering flowers or sprinkler systems. Some participants also considered the frequency of errors when deciding on opt-in or opt-out systems:

> *"I assume that the system acts correctly more often than it does mistakes, so opt-out would be much less effort than constantly having to consent to automation. In reality, we could also measure how often it is right or wrong."* (P3)

3) Warning Factors include elements that pertain directly to the nature and context of warnings. Next to danger and urgency, multiple participants cared more about hazards inside the house than about hazards outside, and consequently preferred higher automation for inside ones. The frequency of warnings was also seen as relevant by some participants, with common problems calling for higher automation. Finally, time of day was brought up by multiple participants. A common concern was how automatically turning off lights at night to save energy during power outages could be dangerous. However, another participant also mentioned the advantages of higher automation at night:

> *"I generally prefer [opt-out systems]. I would still feel like I'm in control, and if a danger occurred at night and I am asleep, I would always prefer [opt-out systems]."* (P4)

6 Discussion

In the following, we will discuss the results of our studies in the broader context of our research question: *Which role does automation play in smart home warning systems, and how do situations of varying danger and urgency influence the users' experiences with such systems?* We begin with the potential we see in SHWSs in Section 6.1. This is followed by Section 6.2 on the design of warnings in the smart home context. Section 6.3 then focuses on the role that automated actions play in safety-critical situations and how potential errors impede desires for automation. Finally, we discuss limitations of our studies and ideas for future work in Section 6.4.

6.1 The Potential of Smart Home Warning Systems

We created a prototypical SHWS and evaluated it in two complementary studies. Study I revealed the potential of such systems to enhance domestic safety by evaluating the general prototype, different automation levels, and modalities. Study II assessed unwanted automation and provided deeper insights into our prototype through interviews and questionnaires. By combining public and local warnings and employing automated actions, i.e., safety protocols, SHWSs can act as an additional warning channel and support users in dealing with dangers. In Study I, we saw that participants were generally willing to use smart home features, particularly in warning situations (Mdn = 7), even though many of them had no prior experience with the technology. These results are also reflected in the perceived usability of our prototype, which was consistently high across all automation levels, with AL0 ($M = 82.0$, $SD = 13.1$), AL1 ($M = 83.6$, $SD = 8.4$), and AL2 ($M = 85.9$, $SD = 7.7$) scoring at the

91st, 94th, and 96th percentiles, respectively [13]. All three automation levels seem to have resulted in highly usable SHWS prototypes. Similarly, Study II found that many participants appreciated the concept of SHWSs.

While we do not expect SHWSs to replace existing warning channels, their unique advantages look promising to enhance domestic safety in the future. With smart home technology becoming more ubiquitous over time, integrating additional features becomes increasingly viable and allows for more proactive hazard mitigation.

6.2 Warning Residents of Dangers

Our studies provided valuable insights into the design of warnings in a smart home context. While we followed prior work's suggestion to present warnings based on their criticality [42], our system *directly* mapped different attributes to modalities. Study I showed that warnings of more dangerous scenarios caused more stress, an attribute generally desirable in warning systems. Study II underscores this result, though our interviews also revealed more nuance to it. In particular, many participants did not view intense heat as very dangerous and therefore thought that the warning was *too* intense, reinforcing the importance of matching the intensity of warnings to the *perceived* criticality of hazards. Interestingly, the LED strip attached to the stairs was rated significantly less suitable in AL2 ($Mdn = 6$) than in AL0 or AL1 ($Mdn = 4.5$), an effect that was not present for the light bulbs spread across the living lab. Our interviews provide an explanation for this effect: light strips are more noticeable and warnings were generally seen as less important if the SHWS could deal with hazards on its own:

> *"I would even go so far as to say that if the plant soil is dry, I don't want to be notified at all — I just want the house to take care of it automatically."* (P10)

This shows that how warnings should be presented also depends on the employed automation level. Importantly, the perception of hazards is subjective: most but not all participants viewed intense heat as less dangerous than we had anticipated. Additionally, while 12 participants in Study II preferred being warned too often, four preferred being warned too rarely. This subjectivity calls for the possibility to customize SHWSs, allowing users to adjust their systems to their own needs. While customization is a well-known requirement in warning apps [56] and includes sounds [58] and types of warnings [44], it is even more relevant in SHWSs. Prior research has already shown that users of general smart home technology shape their ecosystems based on various personal factors [91], and in safety-critical contexts, they may wish to further customize colors or connect other IoT devices to alert them in a way that they appreciate. As an uncritical example of a creative use case for IoT devices informing users of events, participants of a co-design workshop ideated an inflatable cat that would inflate and raise to the ceiling to indicate that the owners' cat is in front of their door [9]. While this example describes more of a playful, everyday life situation, allowing users to customize modalities could help better satisfy their needs. This might also be helpful to impaired individuals, such as elderly or blind smart home users, who can benefit a lot from voice-based interaction [57, 61, 69], but may have unique requirements compared to the average user.

6.3 The House That Saves Me? The Roles of Automated Safety Protocols and Errors

A crucial part of our work has been concerned with automation levels in SHWSs. Both of our studies found that users preferred higher automation during warning situations than in everyday life, showing the distinction between general home automation preferences [74, 75] and automation in safety-critical contexts. Most automated actions were seen positively, including the speed and utility of automating windows, fans, and watering the flowers. However, we also found that most participants did not wish for the maximum automation possible. This manifested in Study I through the significantly reduced feeling of control with AL2 compared to AL1 during the intense heat scenario, whose safety protocol many participants disagreed with. Vice versa, participants had significantly higher control over the gas leak with AL2 than AL1, as they agreed with its safety protocol. Therefore, timer-based automation can maintain control when users disagree with automated actions, which was further underscored in Study II: Participants wanted their SHWSs to give them a window of time to become aware of the situation, though higher danger again called for higher automation, i.e., through shorter timers:

> *"I thought 30 seconds was too long, in dangerous situations [...] I would like [opt-out automation], but with shorter timers, for example 10 seconds."* (P8)

While higher danger and urgency called for higher automation, most participants did not want fully autonomous smart homes, citing fear, unease, or frustration about potential errors. As 37 out of 48 participants in Study I did not want their homes to act *completely* autonomously during warning situations (Mdn = 6), our findings relate to prior research on smart home usage intention, which suggests that high automation is not as important as high controllability [6, 93, 94]. Unsurprisingly, we also saw a strong desire for being able to stop or undo safety protocols, including both touch-based and voice-based input devices. Voice commands directed towards smart speakers and smartphone app were preferred by most participants, though buttons, wall-mounted displays, and smartphone voice input were also seen as useful. Interestingly, the ranking varies between stopping and undoing SHWSs' actions. A likely explanation is that stopping a safety protocol may need to happen quickly and is easier to perform through voice commands or pressing a tangible button, while smartphone apps may allow more precise options to undo actions after the fact. While higher automation resulted in lower interruption, it is important to keep in mind that habituation effects [87] could lead to users overly relying on their SHWSs to save them. Fortunately, we found that high dangers still caused moderate levels of interruption even with AL2.

Next to the hazard's danger and urgency, we also found that the optimal automation level depends on various other factors, some of which relate to unwanted automation and false positives. As errors cannot always be avoided, we argue that in addition to danger and urgency, automation levels need to be based on the cost of unwanted automation due to false positives. However, as the hypothetical fire + sprinkler scenario demonstrated (see Figure 7), this consideration can be difficult: some participants wanted their homes to immediately activate sprinklers if a fire was suspected, while others wished for timers of up to 15 minutes to avoid water damage. Additionally, frequency of errors inversely correlates with automation level, while frequency of hazards and users' impairments positively correlate. Other factors, such as living situation or location, add further complexity. Thus, we again point towards customization: allowing users to shape their systems as they envision them [91] is crucial to ensure that SHWSs do not cause more problems than they solve.

6.4 Limitations and Future Work

While our study presents an important step towards understanding automation in SHWSs, it has some limitations, which will be discussed in the following alongside ideas for future work.

(1) The sample of Study I included mostly students and was therefore, on average, rather young and well-educated. Correspondingly, the average ATI was significantly higher than that of German citizens [34]. While tech-savvy users are an important target group for the smart home market, we recognized the bias and recruited a more diverse sample for Study II. Nonetheless, future work should aim to recruit more diverse samples. In particular, elderly users or people with disabilities [101], including those with impaired mobility or vision [69], could benefit from voice-based SHWSs with automated safety protocols.

(2) We conducted our research inside a living lab to investigate the interaction of automation and warning scenarios. As prior HCI research has shown, realism plays a critical role in ensuring ecological and external validity of studies [60], which is why we equipped the environment with common smart home devices and furniture. Despite our efforts to enhance the natural feel of the laboratory environment, we note that a lab study inherently lacks the realism of a field study. This design was chosen for two reasons. First, elaborate prototypes with automated windows and sliding doors cannot easily be set up in real, domestic spaces. Second, there are particularly high ethical concerns when conducting safety-critical studies in the

field. Unwanted automation of a prototype could cause damage or harm participants, which the controlled setting of our lab study avoided. Future work could look to conduct (long-term) field studies to investigate SHWSs under more realistic conditions, but needs to pay specific attention to such ethical concerns.

(3) Our first study investigated three fixed automation levels and four different warning scenarios. As realistic home setups often encompass more nuanced interstage levels, Study II put specific emphasis on opt-in and opt-out systems and asked participants to choose specific timers for different situations, covering more than the initial three automation levels. However, we have not yet examined the effects of adaptive or adaptable automation [84] based on the criticality of the warnings, which future work should investigate.

(4) During Study II, some participants suggested other ideas for safety protocols, such as automated shutters during heat waves. Similar to research collecting creative use cases of general smart home technology [1], future work should collect users' visions of IoT use cases that enhance safety in the context of SHWSs, for example by co-designing systems and assessing which aspects actually need to be customizable.

(5) Some participants wondered which automated action would take priority if multiple hazards attempted to simultaneously control one device, which may be a highly subjective topic. This question was recently investigated in general smart home contexts [96] and could be investigated in future work on SHWSs.

7 Conclusion

Smart home technology has become widespread over the years and provides ways of improving modern living. Customizable trigger-action programming allows even lay users to automate their domestic spaces, leading to advantages in comfort or efficiency. Other devices increase the users' safety and security, such as smart smoke detectors or security cameras. We see an opportunity for SHWSs, i.e., smart home systems that warn of dangers and automate safety protocols, to enhance residents' safety. To assess the role of automation in such systems, we conducted two complementary user studies based on a prototypical SHWS. Study I invited participants into a living lab, which was equipped with a smart speaker, interfaces to warn residents, and actuators to control electronic devices, windows, and doors. As they experienced one of three automation levels and four different warning scenarios, each with a different level of danger and urgency, we collected data on their perception through questionnaires. Study II extended the focus, gathering qualitative data on the perception of our prototype and exploring how errors, i.e., false positives and unwanted automation, affect participants' attitudes. Our results show that smart home automation is not only useful in everyday life, but can also be helpful in warning situations. We found that participants preferred higher automation in more critical situations and seemed open to the idea of using smart home features to control electronic devices, windows, and doors. Usability was consistently high for all evaluated automation levels. However, we also found that some warnings were regarded as too intense and that the employed safety protocol for intense heat was viewed negatively. These results shed light on how different automation levels can influence control, especially if users disagree with the SHWS's actions. Additionally, while we found that higher automation can reduce perceived interruption, most participants wished for timers before actions are executed, allowing them to understand the situation and intervene if needed. Providing an easy way to cancel or undo safety protocols is also a key part of SHWS design and includes voice-based and touch-based interaction. Overall, we conclude that SHWSs have high potential to ensure the safety and security of residents, and are eager to see how the pervasive nature of IoT technology will influence our homes in the future.

Acknowledgments

This work has been funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center [LOEWE/1/12/519/03/05.001(0016)/72] and by the German Federal Ministry of Education and Research and the Hessian Ministry of Higher Education, Research, Science and the Arts within their joint support of the National Research Center for Applied Cybersecurity ATHENE.

References

[1] Tahir Abbas, Vassilis-Javed Khan, Daniel Tetteroo, and Panos Markopoulos. 2018. How Creative is the Crowd in Describing Smart Home Scenarios?. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3170427.3188509

[2] Nada Alhirabi, Stephanie Beaumont, Jose Tomas Llanos, Dulani Meedeniya, Omer Rana, and Charith Perera. 2023. PARROT: Interactive Privacy-Aware Internet of Things Application Design Tool. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 1:1–1:37. doi:10.1145/3580880

[3] Hamood Alqourabah, Amgad Muneer, and Suliman Mohamed Fati. 2021. A smart fire detection system using iot technology with automatic water sprinkler. *International Journal of Electrical and Computer Engineering (IJECE)* 11, 4 (Aug. 2021), 2994–3002. doi:10.11591/ijece.v11i4.pp2994-3002

[4] Istvan Balogh, Kerstina Ohlsson, Gert-Åke Hansson, Tomas Engström, and Staffan Skerfving. 2006. Increasing the degree of automation in a production system: Consequences for the physical workload. *International Journal of Industrial Ergonomics* 36, 4 (April 2006), 353–365. doi:10.1016/j.ergon.2004.09.009

[5] Nazmiye Balta-Ozkan, Rosemary Davidson, Martha Bicket, and Lorraine Whitmarsh. 2013. The development of smart homes market in the UK. *Energy* 60 (Oct. 2013), 361–372. doi:10.1016/j.energy.2013.08.004

[6] Nazmiye Balta-Ozkan, Rosemary Davidson, Martha Bicket, and Lorraine Whitmarsh. 2013. Social barriers to the adoption of smart homes. *Energy Policy* 63 (Dec. 2013), 363–374. doi:10.1016/j.enpol.2013.08.043

[7] Frank Bentley, Chris Luvogt, Max Silverman, Rushani Wirasinghe, Brooke White, and Danielle Lottridge. 2018. Understanding the Long-Term Use of Smart Speaker Assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (Sept. 2018), 1–24. doi:10.1145/3264901

[8] Arne Berger, Aloha Hufana Ambe, Alessandro Soro, Dries De Roeck, and Margot Brereton. 2019. The Stories People Tell About The Home Through IoT Toolkits. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 7–19. doi:10.1145/3322276.3322308

[9] Arne Berger, William Odom, Michael Storz, Andreas Bischof, Albrecht Kurze, and Eva Hornecker. 2019. The Inflatable Cat: Idiosyncratic Ideation of Smart Objects for the Home. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. doi:10.1145/3290605.3300631

[10] A.J. Bernheim Brush, Bongshin Lee, Ratul Mahajan, Sharad Agarwal, and Stefan Saroiu. 2011. Home automation in the wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Vancouver, BC, Canada, 2115 – 2124. doi:10.1145/1978942.1979249

[11] Sourav Kumar Bhoi, Sanjaya Kumar Panda, Biranchi Narayan Padhi, Manash Kumar Swain, Balabhadrah Hembram, Debasish Mishra, Chittaranjan Mallick, Munesh Singh, and Pabitra Mohan Khilar. 2018. FireDS-IoT: A Fire Detection System for Smart Home Based on IoT Data Analytics. In *2018 International Conference on Information Technology (ICIT)*. IEEE, Bhubaneswar, India, 1–5. doi:10.1109/ICIT.2018.00042

[12] John Brooke. 1995. SUS: A quick and dirty usability scale. *Usability Eval. Ind.* 189 (Nov. 1995).

[13] John Brooke. 2013. SUS: A Retrospective. *JUX - The Journal of User Experience* 8, 2 (Feb. 2013), 29–40. https://uxpajournal.org/sus-a-retrospective/

[14] Jared J. Cash. 2009. Alert fatigue. *American Journal of Health-System Pharmacy* 66, 23 (Dec. 2009), 2098–2101. doi:10.2146/ajhp090181

[15] Narae Cha, Auk Kim, Cheul Young Park, Soowon Kang, Mingyu Park, Jae-Gil Lee, Sangsu Lee, and Uichin Lee. 2020. Hello There! Is Now a Good Time to Talk? Opportune Moments for Proactive Interactions with Smart Speakers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (Sept. 2020), 74:1–74:28. doi:10.1145/3411810

[16] Yawen Chi, Jianxiong Nie, Yanjun Wang, and Daniel Delahaye. 2023. A Review of Situational Awareness in Air Traffic Control. *IEEE Access* PP (Jan. 2023), 1–17. doi:10.1109/ACCESS.2023.3336415

[17] Jane Chung, George Demiris, and Hilaire J. Thompson. 2016. Ethical Considerations Regarding the Use of Smart Home Technologies for Older Adults: An Integrative Review. *Annual Review of Nursing Research* 34, 1 (Jan. 2016), 155–181. doi:10.1891/0739-6686.34.155

[18] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales - Jacob Cohen, 1960. *Educational and Psychological Measurement* 20, 1 (1960). https://journals.sagepub.com/doi/10.1177/001316446002000104

[19] Sven Coppers, Davy Vanacken, and Kris Luyten. 2020. FORTNIoT: Intelligible Predictions to Improve User Understanding of Smart Home Behavior. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 4 (2020), 124:1–124:24. doi:10.1145/3432225

[20] Jessamyn Dahmen, Diane J. Cook, Xiaobo Wang, and Wang Honglei. 2017. Smart secure homes: a survey of smart home technologies that sense, assess, and respond to security threats. *Journal of Reliable Intelligent Environments* 3, 2 (Aug. 2017), 83–98. doi:10.1007/s40860-017-0035-0

[21] Jiaxin Dai, Chao Zhang, Dzmitry Aliakseyeu, Samantha Peeters, and Wijnand A Ijsselsteijn. 2023. The Effect of Explanation Design on User Perception of Smart Home Lighting Systems: A Mixed-method Investigation. In *Proceedings of the 2023 CHI Conference on Human*

Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3544548.3581263

[22] Luigi De Russis and Fulvio Corno. 2015. HomeRules: A Tangible End-User Programming Interface for Smart Homes. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. Association for Computing Machinery, New York, NY, USA, 2109–2114. doi:10.1145/2702613.2732795

[23] Luigi De Russis and Alberto Monge Roffarello. 2018. A Debugging Approach for Trigger-Action Programming. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3170427.3188641

[24] Sarah Delgado Rodriguez, Sarah Prange, Christina Vergara Ossenberg, Markus Henkel, Florian Alt, and Karola Marky. 2022. PriKey – Investigating Tangible Privacy Control for Smart Home Inhabitants and Visitors. In *Nordic Human-Computer Interaction Conference (NordiCHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3546155.3546640

[25] Eren Demir, Erdem Köseoğlu, Radosveta Sokullu, and Burhan Şeker. 2017. Smart Home Assistant for Ambient Assisted Living of Elderly People with Dementia. *Procedia Computer Science* 113 (Jan. 2017), 609–614. doi:10.1016/j.procs.2017.08.302

[26] denysdovhan. 2024. Home Assistant helps me survive during missile and drone attacks. Here is how. www.reddit.com/r/homeassistant/comments/1czm73v/home_assistant_helps_me_survive_during_missile/

[27] Audrey Desjardins, Ron Wakkary, and William Odom. 2023. Investigating Genres and Perspectives in HCI Research on the Home. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Seoul, Republic of Korea, 3073 – 3082. doi:10.1145/2702123.2702540

[28] Ali Dorri, Salil S. Kanhere, Raja Jurdak, and Praveen Gauravaram. 2017. Blockchain for IoT security and privacy: The case study of a smart home. In *2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. IEEE, Kona, HI, USA, 618–623. doi:10.1109/PERCOMW.2017.7917634

[29] Denys Dovhan. 2024. Home Assistant Config by Denys Dovhan - War Safety. https://denysdovhan.com/home-assistant-config/config/war/

[30] Allen L. Edwards. 1951. Balanced Latin-Square Designs in Psychological Research. *The American Journal of Psychology* 64, 4 (1951), 598–603. doi:10.2307/1418200

[31] Mica R. Endsley. 1997. Level of Automation: Integrating Humans and Automated Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 41, 1 (Oct. 1997), 200–204. doi:10.1177/107118139704100146

[32] MICA R. ENDSLEY and DAVID B. KABER. 1999. Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics* 42, 3 (March 1999), 462–492. doi:10.1080/001401399185595

[33] Andy Field, Jeremy Miles, and Zoe Field. 2012. *Discovering Statistics Using R*. Sage. https://www.torrossa.com/en/resources/an/4913501

[34] Thomas Franke, Christiane Attig, and Daniel Wessel. 2019. A Personal Resource for Technology Interaction: Development and Validation of the Affinity for Technology Interaction (ATI) Scale. *International Journal of Human–Computer Interaction* 35, 6 (April 2019), 456–467. doi:10.1080/10447318.2018.1456150

[35] Jörgen Frohm. 2008. *Levels of automation in production systems*. Ph. D. Dissertation. Chalmers University of Technology. doi:10.13140/RG.2.1.2797.7447

[36] Jörgen Frohm, Veronica Lindström, Johan Stahre, and Mats Winroth. 2008. Levels of automation in manufacturing. *Ergonomia - an International journal of ergonomics and human factors* 30, 3 (2008), 36. https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-7253

[37] Bundesamt für Bevölkerungsschutz und Katastrophenhilfe. 2024. *Ergebnisse der Umfrage zum Bundesweiten Warntag 2023*. Technical Report. Federal Office of Civil Protection and Disaster Assistance. https://www.bbk.bund.de/SharedDocs/Downloads/DE/Mediathek/Publikationen/Warnung/umfrageergebnisse-bwt-2023_download.pdf?__blob=publicationFile&v=3

[38] Yi Gao, Kaijie Xiao, Fu Li, Weifeng Xu, Jiaming Huang, and DongWei. 2024. ChatIoT: Zero-code Generation of Trigger-action Based IoT Programs. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (Sept. 2024), 1–29. doi:10.1145/3678585

[39] H. Ghayvat, S. Mukhopadhyay, B. Shenjie, A. Chouhan, and W. Chen. 2018. Smart home based ambient assisted living: Recognition of anomaly in the activity of daily living for an elderly living alone. In *2018 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, Houston, TX, USA, 1–5. doi:10.1109/I2MTC.2018.8409885

[40] Rebecca Grier. 2015. How high is high? A metanalysis of NASA TLX global workload scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 59 (Oct. 2015), 1727–1731. doi:10.1177/1541931215591373

[41] Neilly H. Tan, Richmond Y. Wong, Audrey Desjardins, Sean A. Munson, and James Pierce. 2022. Monitoring Pets, Deterring Intruders, and Casually Spying on Neighbors: Everyday Uses of Smart Home Cameras. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–25. doi:10.1145/3491102.3517617

[42] Steffen Haesler, Marc Wendelborn, and Christian Reuter. 2023. Getting the Residents' Attention: The Perception of Warning Channels in Smart Home Warning Systems. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1114–1127. doi:10.1145/3563657.3596076

[43] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*, Peter A. Hancock and Najmedin Meshkati (Eds.). Human Mental Workload, Vol. 52. North-Holland, 139–183. doi:10.1016/S0166-4115(08)62386-9

[44] Jasmin Haunschild, Marc-André Kaufhold, and Christian Reuter. 2022. Perceptions and Use of Warning Apps – Did Recent Crises Lead to Changes in Germany?. In *Proceedings of Mensch und Computer 2022 (MuC '22)*. Association for Computing Machinery, New York, NY, USA, 25–40. doi:10.1145/3543758.3543770

[45] Andrin Hauri, Kevin Kohler, and Benjamin Scharte. 2022. *A Comparative Assessment of Mobile Device-Based Multi-Hazard Warnings: Saving Lives through Public Alerts in Europe.* Report. ETH Zurich. doi:10.3929/ethz-b-000533908

[46] Tianzhi He and Farrokh Jazizadeh. 2022. Proactive Smart Home Assistants for Automation—User Characteristic-Based Preference Prediction with Machine Learning Techniques. In *Computing in Civil Engineering 2021*. American Society of Civil Engineers, Orlando, Florida, 271–278. doi:10.1061/9780784483893.034

[47] Sascha Henninger, Martin Schneider, and Arne Leitte. 2022. Smart Sirens—Civil Protection in Rural Areas. *Sustainability* 14, 1 (Jan. 2022), 15. doi:10.3390/su14010015

[48] Gaston C. Hillar. 2017. *MQTT Essentials - A Lightweight IoT Protocol: Send and receive messages with the MQTT protocol for your IoT solutions.* https://www.packtpub.com/en-us/product/mqtt-essentials-a-lightweight-iot-protocol-9781787287815

[49] Jannik Hüels and Anne Remke. 2016. Energy Storage in Smart Homes: Grid-Convenience Versus Self-Use and Survivability. In *2016 IEEE 24th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS)*. IEEE, London, UK, 385–390. doi:10.1109/MASCOTS.2016.33

[50] Waheb A. Jabbar, Tee Kok Kian, Roshahliza M. Ramli, Siti Nabila Zubir, Nurthaqifah S. M. Zamrizaman, Mohammed Balfaqih, Vladimir Shepelev, and Soltan Alharbi. 2019. Design and Fabrication of Smart Home With Internet of Things Enabled Automation System. *IEEE Access* 7 (2019), 144059–144074. doi:10.1109/ACCESS.2019.2942846

[51] Andreas Jacobsson, Martin Boldt, and Bengt Carlsson. 2016. A risk analysis of a smart home automation system. *Future Generation Computer Systems* 56 (March 2016), 719–733. doi:10.1016/j.future.2015.09.003

[52] Timo Jakobi, Corinna Ogonowski, Nico Castelli, Gunnar Stevens, and Volker Wulf. 2017. The Catch(es) with Smart Home: Experiences of a Living Lab Field Study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1620–1633. doi:10.1145/3025453.3025799

[53] Timo Jakobi, Gunnar Stevens, Nico Castelli, Corinna Ogonowski, Florian Schaub, Nils Vindice, Dave Randall, Peter Tolmie, and Volker Wulf. 2018. Evolving Needs in IoT Control and Accountability: A Longitudinal Study on Smart Home Intelligibility. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 171:1–171:28. doi:10.1145/3287049

[54] Deepesh Jayasekara, Nai Yeen Gavin Lai, Kok-Hoong Wong, Kulwant Pawar, and Yingdan Zhu. 2022. Level of automation (LOA) in aerospace composite manufacturing: Present status and future directions towards industry 4.0. *Journal of Manufacturing Systems* 62 (Jan. 2022), 44–61. doi:10.1016/j.jmsy.2021.10.015

[55] David B. Kaber and Mica R. Endsley. 2004. The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. *Theoretical Issues in Ergonomics Science* 5, 2 (March 2004), 113–153. doi:10.1080/1463922021000054335

[56] Marc-André Kaufhold, Jasmin Haunschild, and Christian Reuter. 2020. Warning the Public: A Survey on Attitudes, Expectations and Use of Mobile Crisis Apps in Germany. In *Proceedings of the 28th European Conference on Information Systems (ECIS)*. Association for Information Systems, Marrakech, Morocco, 1–16. https://aisel.aisnet.org/ecis2020_rp/84

[57] Sunyoung Kim. 2021. Exploring How Older Adults Use a Smart Speaker–Based Voice Assistant in Their First Interactions: Qualitative Study. *JMIR mHealth and uHealth* 9, 1 (Jan. 2021), e20427. doi:10.2196/20427

[58] Christoph Kotthaus, Thomas Ludwig, and Volkmar Pipek. 2016. Persuasive System Design Analysis of Mobile Warning Apps for Citizens. In *PERSUASIVE*. Salzburg, Austria, 1–12. https://ceur-ws.org/Vol-1817/paper7.pdf

[59] Hermann Leinweber, Jonatan Crystall, Frank Hessel, Florentin Putz, and Matthias Hollick. 2023. Introducing FreeSpeaker – A Modular Smart Home Hub Prototyping Platform. In *Proceedings of the 29th Annual International Conference on Mobile Computing and Networking*. Association for Computing Machinery, New York, NY, USA, 1–3. doi:10.1145/3570361.3614080

[60] Letitia Lew, Truc Nguyen, Solomon Messing, and Sean Westwood. 2011. Of course I wouldn't do that in real life: advancing the arguments for increasing realism in HCI experiments. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems (CHI EA '11)*. Association for Computing Machinery, New York, NY, USA, 419–428. doi:10.1145/1979742.1979621

[61] Brittany Lewis, Tina-Marie Ranalli, Alexandra Gourley, Piriyankan Kirupaharan, and Krishna Venkatasubramanian. 2023. "I... caught a person casing my house... and scared him off:" The Use of Security-Focused Smart Home Devices by People with Disabilities. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3544548.3581007

[62] Xiaoyi Liu, Yingtian Shi, Chun Yu, Cheng Gao, Tianao Yang, Chen Liang, and Yuanchun Shi. 2023. Understanding In-Situ Programming for Smart Home Automation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 2 (June 2023), 1–31. doi:10.1145/3596254

[63] Nikola Marangunić and Andrina Granić. 2015. Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society* 14, 1 (March 2015), 81–95. doi:10.1007/s10209-014-0348-1

[64] Davit Marikyan, Savvas Papagiannidis, and Eleftherios Alamanos. 2019. A systematic review of the smart home literature: A user perspective. *Technological Forecasting and Social Change* 138 (Jan. 2019), 139–154. doi:10.1016/j.techfore.2018.08.015

[65] Karola Marky, Verena Zimmermann, Alina Stöver, Philipp Hoffmann, Kai Kunze, and Max Mühlhäuser. 2020. All in One! User Perceptions on Centralized IoT Privacy Settings. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3334480.3383016

[66] Shahrokh Nikou. 2019. Factors driving the adoption of smart home technology: An empirical assessment. *Telematics and Informatics* 45 (Dec. 2019), 101283. doi:10.1016/j.tele.2019.101283

[67] Linda Onnasch, Christopher D. Wickens, Huiyang Li, and Dietrich Manzey. 2014. Human Performance Consequences of Stages and Levels of Automation: An Integrated Meta-Analysis. *Human Factors* 56, 3 (May 2014), 476–488. doi:10.1177/0018720813501549

[68] Nandita Pattnaik, Shujun Li, and Jason R. C. Nurse. 2023. A Survey of User Perspectives on Security and Privacy in a Home Networking Environment. *ACM Comput. Surv.* 55, 9 (Jan. 2023), 180:1–180:38. doi:10.1145/3558095

[69] Alisha Pradhan, Kanika Mehta, and Leah Findlater. 2018. "Accessibility Came by Accident": Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3173574.3174033

[70] Christian Reuter, Marc-André Kaufhold, Inken Leopold, and Hannah Knipp. 2017. Katwarn, NINA or FEMA? Multi-Method Study on Distribution, Use and Public Views on Crisis Apps. In *Proceedings of the 25th European Conference on Information Systems (ECIS)*. Association for Information Systems, Guimarães, Portugal, 2187–2201.

[71] Christina Rödel, Susanne Stadler, Alexander Meschtscherjakov, and Manfred Tscheligi. 2014. Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '14)*. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/2667317.2667330

[72] Qusay Idrees Sarhan. 2020. Arduino Based Smart Home Warning System. In *2020 IEEE 6th International Conference on Control Science and Systems Engineering (ICCSSE)*. IEEE, Beijing, China, 201–206. doi:10.1109/ICCSSE50399.2020.9171939

[73] Qusay I. Sarhan. 2020. Systematic Survey on Smart Home Safety and Security Systems Using the Arduino Platform. *IEEE Access* 8 (2020), 128362–128384. doi:10.1109/ACCESS.2020.3008610

[74] Eva-Maria Schomakers, Hannah Biermann, and Martina Ziefle. 2020. Understanding Privacy and Trust in Smart Home Environments. In *HCI for Cybersecurity, Privacy and Trust (Lecture Notes in Computer Science)*, Abbas Moallem (Ed.). Springer International Publishing, Cham, 513–532. doi:10.1007/978-3-030-50309-3_34

[75] Eva-Maria Schomakers, Hannah Biermann, and Martina Ziefle. 2021. Users' Preferences for Smart Home Automation – Investigating Aspects of Privacy and Trust. *Telematics and Informatics* 64 (Nov. 2021), 101689. doi:10.1016/j.tele.2021.101689

[76] Alex Sciuto, Arnita Saini, Jodi Forlizzi, and Jason I. Hong. 2018. "Hey Alexa, What's Up?". In *Proceedings of the 2018 Designing Interactive Systems Conference*. Association for Computing Machinery, Hong Kong, 857–868. doi:10.1145/3196709.3196772

[77] Statista. 2024. smart home number of users 2019-2028. https://www.statista.com/forecasts/887613/number-of-smart-homes-in-the-smart-home-market-in-the-world

[78] Statista. 2024. smart home penetration rate 2019-2028. https://www.statista.com/forecasts/887636/penetration-rate-of-smart-homes-in-the-world

[79] Dyah Sugandini, Purwoko, Argo Pambudi, S. Resmi, Reniati Reniati, Muafi Uii, and R.A. Kusumawati. 2018. The role of uncertainty, perceived ease of use, and perceived usefulness towards the technology adoption. *International Journal of Civil Engineering and Technology* 9 (April 2018), 660–669.

[80] S. Tanwar, P. Patel, K. Patel, S. Tyagi, N. Kumar, and M. S. Obaidat. 2017. An advanced Internet of Thing based Security Alert System for Smart Home. In *2017 International Conference on Computer, Information and Telecommunication Systems (CITS)*. IEEE, Dalian, China, 25–29. doi:10.1109/CITS.2017.8035326

[81] Pedro Teixidó, Juan Antonio Gómez-Galán, Fernando Gómez-Bravo, Trinidad Sánchez-Rodríguez, Javier Alcina, and Juan Aponte. 2018. Low-Power Low-Cost Wireless Flood Sensor for Smart Home Systems. *Sensors* 18, 11 (Nov. 2018), 3817. doi:10.3390/s18113817

[82] Federico Trucchia, Hemal Ranawaka Arachchige Dias, Takaharu Suzuki, and Georgia Mackenzie. 2023. Beyond 5G: Envisioning the Future of Mobility in Japan Through Design Fiction. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. ACM, Pittsburgh PA USA, 2109–2122. doi:10.1145/3563657.3596115

[83] Blase Ur, Elyse McManus, Melwyn Pak Yong Ho, and Michael L. Littman. 2014. Practical trigger-action programming in the smart home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 803–812. doi:10.1145/2556288.2557420

[84] Marialena Vagia, Aksel A. Transeth, and Sigurd A. Fjerdingen. 2016. A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Applied Ergonomics* 53 (March 2016), 190–202. doi:10.1016/j.apergo.2015.09.013

[85] Vishakha D. Vaidya and Pinki Vishwakarma. 2018. A Comparative Analysis on Smart Home System to Control, Monitor and Secure Home, Based on Technologies like GSM, IOT, Bluetooth and PIC Microcontroller with ZigBee Modulation. In *2018 International Conference on Smart City and Emerging Technology (ICSCET)*. IEEE, Mumbai, India, 1–4. doi:10.1109/ICSCET.2018.8537381

[86] Thomas van de Werff, Karin Niemantsverdriet, Harm van Essen, and Berry Eggen. 2017. Evaluating Interface Characteristics for Shared Lighting Systems in the Office Environment. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. Association for Computing Machinery, Edinburgh, United Kingdom, 1–12. doi:10.1145/3064663.3064749

[87] Anthony Vance, Brock Kirwan, Daniel Bjornn, Jeffrey Jenkins, and Bonnie Brinton Anderson. 2017. What Do We Really Know about How Habituation to Warnings Occurs Over Time? A Longitudinal fMRI Study of Habituation and Polymorphic Warnings. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2215–2227. doi:10.1145/3025453.3025896

[88] Sarah Theres Völkel, Daniel Buschek, Malin Eiband, Benjamin R. Cowan, and Heinrich Hussmann. 2021. Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3411764.3445536

[89] Jing Wei, Benjamin Tag, Johanne R Trippas, Tilman Dingler, and Vassilis Kostakos. 2022. What Could Possibly Go Wrong When Interacting with Proactive Smart Speakers? A Case Study Using an ESM Application. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22)*. Association for Computing Machinery, New York, NY, USA, 1–15. doi:10.1145/3491102.3517432

[90] Maximiliane Windl, Albrecht Schmidt, and Sebastian S. Feger. 2023. Investigating Tangible Privacy-Preserving Mechanisms for Future Smart Homes. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3544548.3581167

[91] Mikołaj P. Woźniak, Sarah Vöge, Ronja Krüger, Heiko Müller, Marion Koelle, and Susanne Boll. 2023. Inhabiting Interconnected Spaces: How Users Shape and Appropriate Their Smart Home Ecosystems. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Hamburg Germany, 1–18. doi:10.1145/3544548.3581497

[92] Xiao Xue, Xinyang Li, Boyang Jia, Jiachen Du, and Xinyi Fu. 2024. Who Should Hold Control? Rethinking Empowerment in Home Automation among Cohabitants through the Lens of Co-Design. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. Association for Computing Machinery, New York, NY, USA, 1–19. doi:10.1145/3613904.3642866

[93] Heetae Yang, Hwansoo Lee, and Hangjung Zo. 2017. User acceptance of smart home services: an extension of the theory of planned behavior. *Industrial Management & Data Systems* 117, 1 (Jan. 2017), 68–89. doi:10.1108/IMDS-01-2016-0017

[94] Heetae Yang, Wonji Lee, and Hwansoo Lee. 2018. IoT Smart Home Adoption: The Importance of Proper Level Automation. *Journal of Sensors* 2018 (May 2018), e6464036. doi:10.1155/2018/6464036

[95] Yuan Yao, Li Huang, Yi He, Zhijun Ma, Xuhai Xu, and Haipeng Mi. 2023. Reviewing and Reflecting on Smart Home Research from the Human-Centered Perspective. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Hamburg Germany, 1–21. doi:10.1145/3544548.3580842

[96] Ali Zaidi, Rui Yang, Vinay Koshy, Camille Cobb, Indranil Gupta, and Karrie Karahalios. 2023. A User-Centric Evaluation of Smart Home Resolution Approaches for Conflicts Between Routines. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 1 (2023), 45:1–45:35. doi:10.1145/3581997

[97] Nima Zargham, Leon Reicherts, Michael Bonfert, Sarah Theres Voelkel, Johannes Schoening, Rainer Malaka, and Yvonne Rogers. 2022. Understanding Circumstances for Desirable Proactive Behaviour of Voice Assistants: The Proactivity Dilemma. In *Proceedings of the 4th Conference on Conversational User Interfaces (CUI '22)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3543829.3543834

[98] Yanni Zhai and Xiaodong Cheng. 2011. Design of smart home remote monitoring system based on embedded system. In *2011 IEEE 2nd International Conference on Computing, Control and Industrial Engineering*, Vol. 2. IEEE, Wuhan, China, 41–44. doi:10.1109/CCIENG.2011.6008062

[99] Lefan Zhang, Cyrus Zhou, Michael L Littman, Blase Ur, and Shan Lu. 2023. Helping Users Debug Trigger-Action Programs. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 4 (Jan. 2023), 1–32. doi:10.1145/3569506

[100] Serena Zheng, Noah Apthorpe, Marshini Chetty, and Nick Feamster. 2018. User Perceptions of Smart Home IoT Privacy. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 200:1–200:20. doi:10.1145/3274469

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

A Contrasts

In order to break down main and interaction effects of our multilevel linear models, we used the following planned contrasts:

(1) Automation level: For the automation level, we saw AL1 as a baseline condition that requires some user interaction, but also executes some actions automatically. As previous research on smart homes in general shows, users seem to prefer semi-automatic systems [74, 75]. Therefore, we compared our semi-automatic condition with the other two conditions, resulting in the following two contrasts:
   - Contrast 1 (C1) compares low automation (AL0) to medium automation (AL1).
   - Contrast 2 (C2) compares high automation (AL2) to medium automation (AL1).

(2) Warning scenario: For the warning scenario, we were interested in the effects of danger and urgency on the participants' perception. Therefore, we chose the following three contrasts.
   - Contrast 3 (C3) compares the more dangerous scenarios (gas leak and intense heat) to the less dangerous ones (power outage and dry soil).
   - Contrast 4 (C4) compares the more urgent gas leak to the less urgent intense heat scenario.
   - Contrast 5 (C5) compares the more urgent power outage to the less urgent dry soil scenario.

(3) Warning scenario × automation level: For the interaction of our independent variables, contrasts were simply all combinations of the above contrasts.
   - Contrast 6 (C6) compares AL0 to AL1 while comparing the dangerous scenarios to the less dangerous ones.
   - Contrast 7 (C7) compares AL0 to AL1 while comparing the gas leak to the intense heat scenario.
   - Contrast 8 (C8) compares AL0 to AL1 while comparing the power outage to the dry soil scenario.
   - Contrast 9 (C9) compares AL2 to AL1 while comparing the dangerous scenarios to the less dangerous ones.
   - Contrast 10 (C10) compares AL2 to AL1 while comparing the gas leak to the intense heat scenario.
   - Contrast 11 (C11) compares AL2 to AL1 while comparing the power outage to the dry soil scenario.

B Questionnaires

B.1 Study I: Interaction with SHWSs

The following six questions are part of the TLX scale [43] and were asked after the distraction task (sorting cubes), rated on a scale from 1 to 20.

- How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc)? Was the task easy or demanding, simple or complex, exacting or forgiving?
- How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?
- How much time pressure did you feel due to the rate of pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?
- How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
- How hard did you have to work (mentally and physically) to accomplish your level of performance?
- How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

The next questions were asked after each warning scenario, rated on a scale from 1 to 7.

- Did you experience stress due to the system's behavior during the alert?
- Did you experience stress due to the system's behavior after the alert, during the execution?

- To what extent were you interrupted in the execution of your task (sorting) by the alert?
- I had the situation under control.
- Did the system provide you with the assistance you needed in the situation?
- Did the system cause you any negative feelings?
- I would have performed the same actions without the system.

The following questions were asked once after all four scenarios had been completed, rated on a scale from 1 to 7.

- What level of automation should the system in your home have in general?
- What level of automation should the system in your home have during warning scenarios?
- If these functions were available in your home, would you activate / use them in general?
- If these functions were available in your home, would you activate / use them during warning scenarios?
- How suitable do you find the following interfaces of the smart speaker? (sounds, LED lights, button, voice output)
- How suitable do you find the following interfaces of the house overall? (windows, door, monitor, island mode, ventilation, LED strip, light bulbs, smart speaker)

Additionally, the SUS [12] and ATI [34] scales were administered on their respective scales.
SUS:

- I think that I would like to use this system frequently.
- I found the system unnecessarily complex.
- I thought the system was easy to use.
- I think that I would need the support of a technical person to be able to use this system.
- I found the various functions in this system were well integrated.
- I thought there was too much inconsistency in this system.
- I would imagine that most people would learn to use this system very quickly.
- I found the system very cumbersome to use.
- I felt very confident using the system.
- I needed to learn a lot of things before I could get going with this system.

ATI:

- I like to occupy myself in greater detail with technical systems.
- I like testing the functions of new technical systems.
- I predominantly deal with technical systems because I have to.
- When I have a new technical system in front of me, I try it out intensively.
- I enjoy spending time becoming acquainted with a new technical system.
- It is enough for me that a technical system works; I don't care how or why.
- I try to understand how a technical system exactly works.
- It is enough for me to know the basic functions of a technical system.
- I try to make full use of the capabilities of a technical system.

Finally, general questions were asked:

- How familiar were you with the term "smart home" before the study? (1 = never heard of, 5 = very familiar)
- How much experience did you have with smart home technologies before the study? (1 = never been in contact with a smart home, 5 = use them a lot)
- Do you use warning apps?
- What is your age?
- What is your gender?

- If you are a student, what is your major?
- What is your current highest level of education?

B.2 Study II: Mixed-Methods Evaluation of Prototype, False Positives, and Unwanted Automation

- If you had to choose, would you prefer a smart home system that warns you too often or too rarely?
- If you had to choose, would you prefer a smart home system that executes automated actions too often or too rarely?
- How often could you accept the following errors at most? (erroneous warning, unwanted automation of... windows, sprinklers, roller shutters, emergency service notification, lights, entrance door lock, watering flowers, ventilation systems) - answer options included: once a day, once a week, once a month, once a year, less frequently or never
- How much would you like to use the following devices to **stop** automated actions of your house before or as they happen? (smart speaker button, smart speaker voice input, smartphone voice input, smartwatch voice input, wall-mounted displays, smartphone app, smartwatch app, other) - rated on a scale from 1 = would not use at all to 7 = would really like to use
- How much would you like to use the following devices to **undo** automated actions of your house after they happened? (smart speaker button, smart speaker voice input, smartphone voice input, smartwatch voice input, wall-mounted displays, smartphone app, smartwatch app, other) - rated on a scale from 1 = would not use at all to 7 = would really like to use
- (fire + sprinkler) How long should your smart home wait before automatically activating the sprinklers? - rated via free text field and on a scale from 1 = never trigger without consent to 7 = always trigger immediately
- (fire + windows) How long should your smart home wait before automatically closing the windows? - rated via free text field and on a scale from 1 = never trigger without consent to 7 = always trigger immediately
- (poor air quality + purchasing new air filter) How long should your smart home wait before automatically purchasing a new air filter? - rated via free text field and on a scale from 1 = never trigger without consent to 7 = always trigger immediately
- (poor air quality + activating ventilation) How long should your smart home wait before automatically activating the air conditioning? - rated via free text field and on a scale from 1 = never trigger without consent to 7 = always trigger immediately

Additionally, general questions about participants' age, gender etc. were asked (similar to the above).

C Interview Questions

C.1 Perception of Prototype

- Q1 (4x, once per warning scenario): What is your first impression of how the house warned you? Did you like any of them? Did you dislike any of them? Would you have preferred anything else?
- Q2 (4x, once per warning scenario): What is your first impression of the actions? Did you like any of them? Did you dislike any of them? Would you have preferred anything else?

C.2 Unwanted Automation and False Positives

- Q3: In your opinion, what does the optimal level of automation depend on? Are there situations where you would like higher or lower automation?
- Q4: How would you feel if your smart home executed actions trying to protect you, even though the danger is not real / may not want these actions to occur?
- Q5: Out of the two options [opt-in] and [opt-out], which would you prefer and why?

- Q6: Should the level of automation in your home depend on the danger level of a warning?
- Q7: Should the level of automation in your home depend on the urgency of a warning?
- Q8: Should the level of automation in your home depend on the consequences of the automated action?
- Q9: How do you envision your interaction with your smart home if it has executed or is about to execute an action that you do not want to happen?
- Q10: Through which devices would you like to stop or undo actions of your home?

D Codebook

Table 3. Final codebook used to analyze interview data.

(a) Participants' Impressions of Prototype	
Warning	⟨positive⟩ ⟨negative⟩ ⟨LED bulbs⟩ ⟨LED strip⟩ ⟨speaker sound⟩ ⟨speaker message⟩ ⟨other suggestion⟩
Execution	⟨positive⟩ ⟨negative⟩ ⟨windows⟩ ⟨sliding door⟩ ⟨fan⟩ ⟨ceiling lights⟩ ⟨watering flowers⟩ ⟨fan⟩ ⟨other suggestion⟩ ⟨utility⟩ ⟨speed⟩
Interaction	⟨positive⟩ ⟨negative⟩ ⟨control⟩ ⟨modality⟩
(b) Oversensitive Smart Home Warning Systems	
Errors	⟨unpleasant⟩ ⟨safety⟩ ⟨trust⟩ ⟨frustration⟩ ⟨fear⟩ ⟨control⟩
Interaction	⟨cancel⟩ ⟨reverse⟩ ⟨settings⟩ ⟨timing⟩ ⟨voice-based⟩ ⟨touch-based⟩
(c) Factors Influencing Automation Level	
Personal and Context	⟨activity⟩ ⟨age⟩ ⟨instructions⟩ ⟨impairments⟩ ⟨living situation⟩ ⟨safety attitude⟩ ⟨ATI⟩ ⟨location⟩
System Design	⟨frequency of errors⟩ ⟨reversibility⟩ ⟨damage or cost⟩ ⟨utility⟩ ⟨effort of manual execution⟩
Warning	⟨danger level⟩ ⟨urgency⟩ ⟨location of hazard⟩ ⟨frequency⟩ ⟨time of day⟩