



Article

Invisible but Understandable: In Search of the Sweet Spot between Technology Invisibility and Transparency in Smart Spaces and Beyond

Sarah Diefenbach ^{1,*} , Lara Christoforakos ¹, Daniel Ullrich ² and Andreas Butz ²

¹ Department of Psychology, Ludwig-Maximilians-Universität München, 80539 München, Germany

² Department of Computer Science, Ludwig-Maximilians-Universität München, 80539 München, Germany

* Correspondence: sarah.diefenbach@psy.lmu.de; Tel.: +49-89-21805202

Abstract: Smart technology is already present in many areas of everyday life. People rely on algorithms in crucial life domains such as finance and healthcare, and the smart car promises a more relaxed driving experience—all the while, the technology recedes further into the background. The smarter the technology, the more intransparent it tends to become. Users no longer understand how the technology works, what its limits are, and what consequences regarding autonomy and privacy emerge. Both extremes, total invisibility and total transparency, come with specific challenges and do not form reasonable design goals. This research explores the potential tension between smart and invisible versus transparent and understandable technology. We discuss related theories from the fields of explainable AI (XAI) as well as trust psychology, and then introduce transparency in smart spaces as a special field of application. A case study explores specific challenges and design approaches through the example of a so-called room intelligence (RI), i.e., a special kind of smart living room. We conclude with research perspectives for more general design approaches and implications for future research.

Keywords: transparency; invisibility; explainable AI; trust; smart technology; smart spaces



Citation: Diefenbach, S.; Christoforakos, L.; Ullrich, D.; Butz, A. Invisible but Understandable: In Search of the Sweet Spot between Technology Invisibility and Transparency in Smart Spaces and Beyond. *Multimodal Technol. Interact.* **2022**, *6*, 95. <https://doi.org/10.3390/mti6100095>

Academic Editors: Martin Tomitsch, Georgios N. Yannakakis and Cristina Portalés Ricart

Received: 10 August 2022

Accepted: 18 October 2022

Published: 20 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. The “Church of Invisibility”

Hello Traveler, let me introduce you to the Church of Invisibility. We disciples of the church value invisibility more than anything else—the void is precious, as we say with a smile. Following this premise, everything technical must step back behind the curtain and only the mere necessities shall remain visible.

Products made in line with the tenets of the “Church of Invisibility” are generally highly appreciated. This is no wonder, since our attention is limited and anything that helps us simplify life and preserve our energy is a good thing.

That is, unless it comes to one of those unfavorable events, like the moose incident. It happened after Jill bought her new car, a fantastic piece of art stuffed with technology—it could even drive in autopilot if you wanted it to. In addition, it included a number of extra features, such as the “eye of the beholder” (a catchy phrase invented by marketing people)—a night vision projected onto the windshield where potential hazards were highlighted with surrounding red boxes. Seemingly the best part? You did not get annoyed by technology. All those wonderful features remained hidden and silent but were ready when needed; there was nothing that made you care about their exact functioning and capabilities.

One night, Jill was on her way home, her ride aided by autopilot and the eye of the beholder. When the moose appeared in the middle of the road, Jill saw it. The eye of the beholder saw it; And Jill saw the eye seeing it. In the moment right before impact, Jill asked herself how long the car would wait before engaging evasive maneuvers.

1.2. The “Church of Transparency”

Hello Traveler, don't be misguided by the invisibility apologists. We disciples of the Church of Transparency know that only visible things carry value—information leads to wisdom, as we say with a smile. Following this premise, technology should involve the user in all its decisions and actions, confront the user with all available options and their consequences, and thereby establish full transparency at all times.

Products following the guidelines of the church of transparency are scarce these days. In the past, transparency was often naturally given, e.g., one could literally see how a typewriter worked. Nowadays, technological artefacts are increasing and getting more complicated at the same time. Fewer and fewer people aspire to exactly understand the technology (and its limits), to spend time with manuals and tutorials, and to adjust the settings before first usage. To some, however, it is a necessity: knowing how a product works, what it is capable of, and what happens with any collected private data. These individuals want a sophisticated and autonomous use of technology, instead of blindly trusting in an algorithm. Devoted users love those products—until the pursuit of transparency and control go too far even for them.

It was a hot summer day when Jack was on his way home in his new car. He was tired and annoyed by the traffic jam, which would last for at least another hour, while his son was at home waiting for him (and his birthday present). An ideal situation to engage the super-smart navigation system, which could surely find a quicker route. The catch? A modern GPS is an incredibly complex system, which Jack could not use until the setup wizard had discussed Jack's preferences in every detail. Jack had to specify his typical fuel consumption, preferred roads and typical speeds depending on different weather conditions, and so on and on. In addition, Jack had to watch the mandatory tutorial video before he could unlock the navigation feature—to make sure he really understood how the navigation worked, why a particular road was recommended to him, and how changing his settings would affect the recommendations. What a pity to have missed the opportunity to watch the video beforehand.

1.3. The Tension between Smart and Invisible versus Transparent and Understandable

Today, smart technology is present in many areas of everyday life. It takes many tasks off humans' hands and adapts to users' needs and wishes, presumably better than the user could ask for. For example, an intelligent sleep app, e.g., [1,2] can tell us when it is time to go to bed and calculates the biologically ideal time to wake up again. A smart home system adjusts the light atmosphere to our current mood and tasks of the day, e.g., [3,4]. Similarly, we rely on algorithms and their advice in crucial life domains, such as finance [5] and healthcare [6]. In general, the exact functioning of these technologies remains invisible to us as it is occurring “behind the scenes”. We do not have to comprehend how or why the sleep app arrives at the conclusion to send us to bed at 11 p.m. As long as it works and we feel recovered when we start the day, it still makes our life easier, and that is what appears to matter most. Similarly, despite not being fully autonomous yet, many cars currently on the market are more like computers than vehicles in a classical sense. The car itself automatically performs smart adjustments like engaging personalized profiles for different drivers and many other tasks. Yet, even the smartest car sometimes has problems and the smarter the car, the more difficult it can be to diagnose the problem. Instead of a quick look under the hood to see what is wrong, the car generates a mystical error code and we are prompted to consult a “specialist”. Similarly, with many other devices (e.g., Apple smartphones or computers, see [7] for an example), the user is not meant to explore, understand, unscrew or fix them—the technology remains, literally, behind the scenes. While on the one hand, invisible technology is a good thing because it reduces complexity and relieves the user from a mental burden, on the other hand, it also creates a certain level of opacity and with it insecurity and lack of autonomy. Thus, the smarter the technology, the less transparent it becomes.

From a user experience (UX) perspective, both extremes, total invisibility and total transparency, come with specific problems and thus do not form reasonable design goals. As illustrated by the introductory examples and the two opposing “churches”, total transparency tends to be too demanding and confusing. It can come with cognitive overload, thereby negatively affecting the overall UX. Consequently, many people might refrain from using a technology that openly presents its full functionality and complexity, since the many options and decisions become simply too overwhelming. They might also be intimidated by being made aware of the sheer amount of technology around them, which they otherwise might not notice or care for. For this reason, invisible technology can be more inviting for many people. It makes their life easier in an unobtrusive manner and, ideally, they do not even realize it is there. The technically uninterested user is happy as long as it works. The significant downside is that an inappropriate mental model of what the technology does and is capable of can also have negative effects, such as a feeling of insecurity, disappointment, mistrust, but also overtrust, e.g., [8]. This research explores the potential tension between smart and invisible versus transparent and understandable technology.

We discuss related theories in the fields of explainable AI (XAI) and trust psychology, and then draft specific challenges and design approaches of transparency in smart spaces. This is further illustrated by the example of a so-called room intelligence (RI), i.e., a special kind of smart living room. After a brief discussion of related concepts in human–computer interaction (HCI), we present the results of an exploratory workshop to better understand the requirements of potential users of the RI and assess the relevance of explainability and transparency. We conclude with research perspectives for more general design approaches and an outlook on future research questions.

2. Related Research: Explainable AI and Interpersonal Transparency

Transparency is also one of the core elements of the currently booming field of explainable AI (see [9]). Explainable AI communicates to the user what intelligent technology does and how it arrives at particular conclusions e.g., [6,10,11]. Explainable AI aims at enhancing trust in the technology but also providing opportunity to intervene or improve the algorithm at work. Typically, this means the transparency of algorithms and explaining the predictions of AI systems, as in, explaining the reasons why a rule-based expert system rejects a credit card payment or why a shopping website recommends a certain product. It could also mean making transparent the particular pixels by which a picture is classified as a particular object, e.g., a horse and not a zebra [11]. For instance, problems can occur when AI learns based on spurious correlations in the training data [10,12]. Samek and Müller [10] found that an algorithm for picture classification was often not detecting the object of interest, but was utilizing correlations or context in the data. Boats would be recognized by the presence of water or trains by the presence of rails. In other instances, the system responded to spurious contexts such as a copyright watermark that was present in many horse images in the training data [10]. In such cases, transparent AI that lays open on which pixel images the classification was based could help to detect such faulty algorithms. Deeper explanations of AI recommendations promise to be interesting to particular groups of recipients. Samek and Müller [10] discuss the example of an AI application in the medical domain. Here, the AI system could provide simple explanations to the patients (e.g., too high blood sugar), more elaborate explanations to the medical personnel (e.g., unusual relation between different blood sugar parameters) and aggregated explanations (e.g., patterns the AI system learnt after analyzing many patients) to institutions such as hospitals.

Besides referring to adequate technical understanding, transparency also has a psychological component. In this case it relates to trust in the technology and its perception as “a social counterpart”. In the same way that we prefer interactions with people we trust and understand, we might also prefer interactions with technology that we trust and understand. Considering interaction with technology in analogy to interpersonal interaction might help in achieving an appropriate level of transparency in smart systems,

e.g., [13]. In interpersonal interaction, “understanding” someone, does not mean to have insight into its “full functioning” (. . . most people probably do not have this themselves). It rather implies experiencing someone as a reliable counterpart without hidden motives. Unless they acted in a consistent way and communicated their thoughts and intentions openly, we would not find our counterparts’ behavior trustworthy. Similarly, organizational transparency and openness are acknowledged as key components of trust in the public relations literature [14,15], and “communication visibility” has been identified as a predictor of trust among coworkers [16]. In human collaboration, it was observed that visibility and transparency of actions improved collaboration between human actors [17].

In psychological literature, interpersonal trust consists of multiple components, such as cognition-based trust, depending on rational thoughts based on facts, and affect-based trust, based on the emotional connection of a relationship, e.g., [18]. Regarding the question of what builds trust, many models highlight competence and warmth/benevolence [19,20]. In other words, in order to trust someone (to fulfill a particular task for us, e.g., a doctor suggesting the right medicine), we must be convinced that the other is capable of handling the task and that the other means well. Similarly, in human–technology interaction, both components of trust emerged as generally relevant, e.g., [21]. Especially cognitive-based trust and the competence component seem to depend on transparency, i.e., how well the system’s functionality and capabilities are communicated.

3. Transparency in Smart Spaces

While the field of XAI is typically concerned with the transparency of algorithms and explaining the predictions of AI systems on a technical level, the presence of AI in our physical environment, such as the smart home context, adds a different dimension to the need for transparency. Naturally, issues such as risk management and privacy bear special relevance in this context. When technology is evidently present or even ubiquitous, and especially when it is able to impinge on the user within this context, it becomes critical how this exertion of influence can be designed to appear sensible, trustworthy, transparent, and non-intrusive. For example, overtrust in technology due to lacking transparency of the AI’s actual abilities could lead to severe damages to person and property if the user left the flat while the intelligent kitchen was preparing a meal. Furthermore, smart environments are not exclusively used by one person, but transparency also concerns potential visitors. For example, a visitor might be impressed and upset at the same time when the room intelligence hands them their coat after they mentioned they would have to leave in a minute, because they did not expect that “someone else” was listening to their conversation.

Thus, in addition to transparency in the sense of comprehensibility (understanding why something is happening or why something is recommended), transparency in smart spaces also relates to the communication of the mere presence or activation of technology (understanding that technology is at work, there is more happening than meets the eye). This relates closely to the early concept of “intelligibility”, as introduced in the field of context-aware systems by Belotti and Edwards [22]. As argued, “context-aware systems that seek to act upon what they infer about the context must be able to represent to their users what they know, how they know it, and what they are doing about it” ([22] p. 201). Only then can users make informed decisions based on context. However, the proposed design principles of intelligibility are mainly discussed in the context of graphical user interfaces and it still needs to be explored to what degree they can be transferred to achieve transparency in smart spaces.

3.1. The Room Intelligence—A Special Kind of Smart Living Room

In order to better understand the questions related to transparency in smart spaces we examined the example of a smart living room. As a concept for exploration, we used the concept of the “Room Intelligence” (RI) as suggested by Diefenbach and colleagues [23]: Considering the many tasks that technology can take care of, smart home environments typically consist of a diverse range of devices, each with different UI concepts. In order to

manage and, to a certain extent, hide the inherent complexity and heterogeneity, it seems promising to represent the environment as a coherent whole, rather than the sum of its parts. In order to achieve this, the RI proposed by Diefenbach and colleagues [23] presents the environment as an omnipresent intelligence and coherent entity, integrating various smart devices (e.g., lighting, sound, a robot arm that can serve drinks and support in the kitchen) in a holistic interaction design, blurring the line between the physical and the digital world. One might consider this in analogy to a smart car, which integrates many separate devices (e.g., navigation system, climate control, driving assistance), yet is perceived as a coherent entity. The RI does not appear as an additional element in the room, but essentially *is* the room. With people, we cannot always tell what exactly is going on in their minds and understand all the reasons behind their actions. Similarly, the overarching UI concept of the RI suggests that many details of what is happening in the environment will remain hidden to the user. Science fiction fans may feel reminded of movies, such as *Star Trek* or *2001: A Space Odyssey*, where the intelligence mostly manifests itself as an omnipresent voice. In fact, invisibility is explicitly mentioned as a central characteristic of the RI, describing the RI as “an assistant in the background” and “acting behind the scenes”, and “someone” who is “there when you need it, but does not dominate your home” ([23] p. 1/2).

Regarding the technical operationalization, the room intelligence might combine different elements and modalities (see also ([23] p. 4)). For example, visual design elements (e.g., lightning, displays) could attract and guide attention, while projective displays could create output in any location, such as on the walls and the floor. Acoustic elements for spatial sounds such as distributed speakers could be used to represent the “voice” of the RI but possibly also other acoustic feedback or services (e.g., playing music). Robotic arms at a ceiling rail system that can move to any position in the room could provide “helping hands”. In addition, the RI might also use schematic representations of the RI’s presence, e.g., a “face” of the RI, realized through screens or 3D representations of the face such as a grid of movable cylinders integrated in the wall.

3.2. Related Concepts in Human–Computer Interaction Research

The basic idea of a coherent UI as a mediator, for multiple connected devices that assists the user with various tasks involving digital and physical objects in the surrounding environment, is related to various concepts and ongoing research in human–computer interaction (HCI). This includes, for instance, ubiquitous computing, tangible interaction, augmented reality (AR) and intelligent assistants in the home context.

The vision of “ubiquitous computing”, originally formulated by Weiser [24], describes a world in which many things possess computing power and interact with each other to provide new functionality. This vision inspired a comprehensive body of research on how to implement environments in which digital and physical elements are blended. For example, the early idea of tangible interaction [25] directly addresses the combination of physical and digital elements in the real world at an interaction level. The field of augmented reality [26] also mixes physical and digital elements, albeit by means of see-through devices, such as handheld or head-mounted displays. Mobile augmented reality (MAR) integrates computer-generated virtual objects with physical environments for mobile devices (e.g., smartphones, head-worn wearables) thereby enabling users to perform seamless transitions from the physical world to a mixed world with digital entities (for a recent review see [27]). The special variation of spatial augmented reality, e.g., [28–30] uses room scale displays, bypassing the need to wear or hold any special equipment.

Similarly, research on digital assistants explored the potential of connecting intelligent assistants into ubiquitous computing environments in an Internet of Things (IoT) context. For example, Santos and colleagues [31] discuss the technical requirements to integrate wireless sensor networks with the Internet properly by considering the heterogeneity of objects as well as the diversity of communication protocols and enabling technologies. Hervás and Bravo [32] study approaches to the management of contextual information in ambient intelligent systems (i.e., to recognize the person in need of information, the

environment, and the available devices and services) in order to provide proactive user interface adaptation. Kim and colleagues [33] study intelligent assistants in the IoT-context and focus on the relevance of visual embodiment. As Kim argues, many current-state commercial assistants used in the home context, such as Amazon's Alexa, mainly focus on voice commands and voice feedback, and lack the ability to provide non-verbal cues, which are an important part of social interaction. Consequently, their study explores how a visual embodiment through AR affects users' confidence and the perceived social presence of the assistant.

Finally, on a higher conceptual level, one might ask questions about the right way of how an intelligent assistant positions itself to the user and the way it cooperates with the user. As argued in a recent review by Dhiman and colleagues, "the assistants of the future will not only have to be trustworthy, respect our privacy, be accountable and fair, but also help us flourish as human beings" ([34] p. 1). This ambition comes with many ensuing questions, which Dhiman and colleagues cluster into questions on functionality (what an intelligent assistant does), outcomes (its benefit for the user), context (entities, contextual conditions and modes of interaction), design trends (how assistants have been designed in the past), and evaluation (metrics for evaluating an intelligent assistant).

Returning to the vision of ubiquitous computing, it can be stated that the technological requirements have already been achieved widely in many of today's smart home technologies (e.g., media appliances, thermostats, wirelessly controlled power sockets). However, on an interaction level, the vision of ubiquitous computing has not yet been achieved. On the contrary, the configuration of and interaction with such a collection of devices has become so complex that it currently rather stands in the way of widespread adoption and use. In line with this, Weiser himself later claimed that a common and understandable interaction paradigm will be one of the key challenges for ubiquitous computing. Until now, the appropriate user interface remains a key question in the design of ambient intelligence and smart environments. As already discussed by Butz ([35] p. 535), "there are many ways to create smart environments and to realize the vision of ambient intelligence. But whatever constitutes this smartness or intelligence has to manifest itself to the human user through the human senses" and "the devices which create these phenomena (e.g., light, sound, force . . .) or sense these actions are the user's contact point with the underlying smartness or intelligence".

Altogether, existing work in HCI still lacks insights about how different design cues and modalities of interaction can be integrated to shape the overall perception of system intelligence and entity, and finally, also the users' perceived own role within such an intelligent environment as the RI (see also ([23] p. 3)).

4. Exploratory Workshop

4.1. Materials and Methods

In order to get a first general evaluation of the idea of the RI and the requirements of potential users we conducted an exploratory workshop with ten researchers in the field of psychology (70% female, 30% male; ranging from 23 to 60 years). Due to restrictions during the COVID-19 pandemic, the workshop was conducted online, using the video conferencing software Zoom. The workshop lasted for about ninety minutes. In the beginning, we presented a slide about the vision of the RI as suggested by Diefenbach and colleagues [23]. More specifically, we used the textual description presented at the beginning of the paper by Diefenbach ([23] p. 1) that embeds the RI vision in the report of a time traveler and a discussion about the way people will be living in the future, emphasizing its unobtrusiveness and invisibility.

The other day I had a visitor: she was a time traveler from the year 2068. For the moment, the reason for her visit is less important. More important is our discussion about the way people will be living in 2068 and the role of technology in their daily life. "I guess it will all look very modern and futuristic, huh? Flying cars, robots serving you breakfast, connected devices and technology everywhere you look around?"—"Oh

no”, she laughed, “thank God it is pretty much the opposite. In fact, you hardly even see technology anymore”, and, with a more nostalgic look in her face, “but you are right, there was a time when technology was all over the place. New devices and zillions of platforms just to manage everyday life. It felt like becoming a slave to it. This was when I was in my 20s, I still remember that horrible time. Then, luckily, things changed. Now we found alternative ways to connect the many devices in a smart way and to integrate them smoothly in a comfortable environment. Officially, we call it a ‘room intelligence’, but I just say ‘roomie’. It is there when you need it, but it does not dominate your home. And it respects the situation: for example, when I am having visitors in my office, the room intelligence brings coffee and cake in a decent and effective way. I can focus on welcoming the guests and on the conversation”. I was stunned by her descriptions: “Wow, interesting! Tell me more about this room intelligence! How do you communicate with it? What does it look like? How do you actually know it is there if you do not see it?”—“Good questions!” she smiled, “I see you got the point.” ([23] p. 1/2)

After this introduction, participants discussed their impression and ideas of the RI along the following guiding questions: Which chances or benefits could the RI offer to you? Which challenges or risks could you imagine? What kind of UI/point of contact would you expect to interact with the RI?

4.2. Results and Discussion

Participants’ contributions relating to the general potential and associated chances with the RI comprised diverse ideas and usage scenarios. For example, many participants mentioned practical support in situations where one might feel the need of a personal clone or a “third arm” (e.g., while cooking). In addition, participants envisioned emotional support through the RI, e.g., improving one’s mood or deescalating conflicts by playing music. Other comments referred to preparing the room and welcoming the user back home, for example, with cozy lighting, an ideal temperature, monitoring the air quality, or preparing a cup of tea. Finally, another cluster of comments referred to hosting guests (e.g., taking their coat, serving coffee), freeing the individual to focus on the guest and the conversation instead of practical tasks.

Comments about the risks and requirements often touched on issues of explainability and transparency, as exemplified by the following paragraphs:

How do I know which parts of the RI are active?

Several participants wished for a central interface that would inform about ongoing processes in the RI. One participant described this need by referring to her experience with using Bluetooth headphones: “(. . .) when I put my Bluetooth headphones on, I receive the information ‘I am connected to three devices’, but not to which devices. And then I know, ok, I am listening to this Zoom call in my headphones but I don’t know: ‘Will they ring if my phone rings? Or if my roommate wants to listen to music, will this go straight into my zoom call?’”.

How do I start—and end—the dialogue with the RI?

Regarding the UI/point of contact with the RI, participants wished for a clearly defined central point of contact in the sense of an on/off button, in order to provide clarity on whether the RI is activated or not. Participants compared this to a car that needs a key to start or a hotel room that is “activated” by a key card. This requirement also referred to potential visitors who should be able to identify that they were being in a smart space and perhaps opt out of it. As one participant stated “(. . .) Somehow, I would find it weird, if I was visiting and knew that my whole conversation was recorded”.

The room knows what’s good for me—but does it really know?

While participants valued the idea of smart adjustments to meet their needs (e.g., when I come home, the RI detects my mood and plays music that is good for me), they also discussed the risks of giving full control to the RI. In particular, they emphasized the importance of an opportunity for manual intervention, in case they were not convinced

of what the RI “thought” was good for them, or in case the RI had “misunderstood” the situation. For example, no essential functions, such as turning on/off the fridge, should be controlled solely through the RI and alternative manual usage should still be possible. Although participants thought that the RI could be a helpful partner in many respects, they still questioned whether it would actually manage to identify the right action in every moment. This caution was rooted in their previous experiences with existing smart home elements. As one participant stated: “There is nothing more annoying than rooms with automatic blinds”. In this respect, participants also mentioned that the right degree of transparency would depend on the specific user. As one participant explained “(. . .) when I think about my mom, it could be possible that less transparency would be better and that she would just need a room personality saying: ‘everything is functioning in interaction, don’t worry, I will handle this for you’”.

Overall, the perceived opportunities, as well as risks and requirements associated with the RI show some parallels to user demands surveyed in the smart home literature. For example, a recent review of benefits and barriers associated with commercial smart home technologies in Europe [36] listed “convenience and controllability” as one of the main benefits. Typical interview statements were “Anything that helps [consumers] to reduce their mental load on tasks” or “anything that makes you more comfortable and easier for you to get the outcome you want without having to consciously think about how to achieve that outcome” ([36] p. 8). This corresponds with our participants’ mentions of practical support (a “third arm”) or tasks like maintaining the ideal temperature or monitoring the air quality. The review by Sovacool and colleagues also mentioned health benefits; however, this primarily referred to the ability to alert relatives or health professionals in case of an emergency or health diagnosis, but not emotional support or caring for a “cozy atmosphere” as suggested by our study participants. Regarding risks and barriers, the smart home literature mentions also the fear of loss of personal control and autonomy [36], and, in turn, a strong interest in continuous control over system states [37], which is reflected by our participants’ mentions of the risks and requirements regarding the RI. However, the discussion of such issues in the smart home literature typically remains on an abstract level, and does not address the specific tension between transparency and invisibility or related design principles.

5. Discussion and Future Work

Based on our findings and reflections above, we see multiple contributions of our study and suggest further research on different levels. In particular, this concerns (1) a general emphasis on the tension between technology invisibility and transparency, and therefore a need for more research and adequate design approaches; (2) The psychological dimensions related to the issue of technology invisibility versus transparency and individual preferences as to the right level of transparency; and (3), more attention to technology transparency beyond the field of screen-applications and algorithms in recommender systems. The following paragraphs discuss these issues in more detail.

Invisible technology is intransparent by nature. While there are comprehensible reasons to push technology to the background and reduce its visibility to the desired effect for the user, this produces new problems, as outlined above. On the other hand, complete revelation of technical details will create transparency, but overexpose the entire complexity. Considering these two opposing poles of a spectrum, future research should explore design approaches hitting the “sweet spot” of calibrated transparency on this spectrum.

From a psychological perspective, the issue of technology invisibility versus transparency also hints at a tension between two human needs. On one side is the desire to understand and control what is happening. On the other side is the desire for comfort, freedom, and autonomy, as well as the wish to delegate a measure of responsibility, in order to concentrate on other things. This tension is present in many areas of life. For example, in the working context, a secretary has the difficult task to find the right level of transparency in communication routines and keeping others in the loop of a conversation.

Asking a manager about everything can be experienced as disrespectful and as stealing time from more important issues. At the same time, excluding managers from certain decisions about seemingly trivial issues (e.g., choosing catering for the team Christmas party) can be experienced as an unwanted loss of control. Obviously, achieving the right level of transparency in human communication is not an easy issue.

The same applies to the sphere of human–technology interaction, where neither absolute transparency, nor total invisibility form ideal paths of communication. For example, many users are annoyed by the many status messages and routine notifications on their computers, but also object to blind transfer of private data or computer updates that are initiated without allowing to postpone them to a more suitable point in time. In fact, dialogue boxes and cryptic error messages may not be the most effective tool to motivate humans to become “informed users” and really understand what is happening behind the scenes. Thus, an important question remains, of how to provide an adequate technical environment for more conscious user decisions.

In our vision, a more human-centered dialogue between technology and user is an important key. For example, when “starting their relationship” (e.g., setting up a computer) there could be a general conversation about the user’s desired degree of transparency and preferences of involvement in different areas. The user may wish to dive deeper and make better, more autonomous decisions in some areas (e.g., privacy issues, what kind of data is being transferred), but prefer to lay back and relax and let the technology decide in other areas (e.g., automatic adjustments of screen brightness when the environment has changed). As highlighted in the above workshop results, there is not likely to be one right level of transparency for all users. Future research should refer to individual preferences and to the question of what degree of invisibility and transparency is adequate for people with different levels of technical understanding.

Finally, we suggest paying more attention to technology transparency beyond the field of screen-applications and algorithms in recommender systems, which XAI has focused on so far. As laid out above, one area of particular interest is the field of smart spaces and to what degree current design principles of XAI are transferable. Consequently, our next steps of research in this area will explore the communicative building blocks for creating transparency and credibility in a smart space. This may cover principles from XAI but also elements of transparency in interpersonal communication (see above).

6. Conclusions

In conclusion, the search of the sweet spot between technology invisibility and transparency turned out to be quite a complex endeavor. What becomes apparent from the research is that the right level of transparency depends on many factors of context and person, and may even change over time. Users may want more transparency with growing expertise, and their desire to know or not know what is happening behind the scenes may vary depending on the context of application. Furthermore, as also discussed in privacy research, people may prefer limited transparency in certain situations (e.g., about dubious app permissions) to prevent cognitive dissonance. If, for example, a messaging app asks for multiple dubious permissions to track data, but not using the app is experienced as social exclusion, people may still use the app but prefer not to look in detail at the data transferred to the provider.

While finding a universal sweet spot seems rather elusive, we can at least look to our vision, as represented by the “Church of Equilibrium”:

Hello Traveler, I am glad that you finally found us. We disciples of the Church of Equilibrium know that following extremes is a sure path to the valley of unhappiness—everything should be perfectly balanced, as we say with a smile. Following this premise, we should take into account the user’s needs and expertise, the usage context, as well as the technology’s complexity and potential to cause harm. Sounds complicated? True, but good design never comes free of charge . . .

Author Contributions: Conceptualization, S.D., L.C., D.U. and A.B.; methodology, S.D. and L.C.; writing—original draft preparation, S.D. and D.U.; writing—review and editing, L.C. and A.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Research Foundation (DFG), Project PerforM (425412993) as part of the Priority Program SPP2199 Scalable Interaction Paradigms for Pervasive Computing Environments.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sleep Cycle: Sleep Recorder. 2022. Available online: <https://play.google.com/store/apps/details?id=com.northcube.sleepcycle&hl=de&gl=US> (accessed on 9 June 2022).
2. Sleep Theory Sleep Tracker. 2022. Available online: <https://play.google.com/store/apps/details?id=com.noxgroup.app.sleeptheory> (accessed on 9 June 2022).
3. Loxone Smart Home. 2022. Available online: <https://www.smarthomenord.de/licht/> (accessed on 9 June 2022).
4. Smart Energy. 2022. Available online: <https://smartenergy-elektrotechnik.de/smart-home-licht/> (accessed on 9 June 2022).
5. MacKenzie, D. Material signals: A historical sociology of high-frequency trading. *Am. J. Sociol.* **2018**, *123*, 1635–1683. [CrossRef]
6. Holzinger, A. From machine learning to explainable AI. In Proceedings of the 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA), Košice, Slovakia, 23–25 August 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 55–66. [CrossRef]
7. Paul, P. Apple's New 12-inch MacBook Closely Guards Its Secrets with Screws, Solder, and Glue. 2015. Available online: <https://www.macworld.com/article/225313/apples-new-12-inch-macbook-closely-guards-its-secrets-with-screws-solder-and-glue.html> (accessed on 9 June 2022).
8. Ullrich, D.; Butz, A.; Diefenbach, S. The development of overtrust: An empirical simulation and psychological analysis in the context of human–robot interaction. *Front. Robot. AI* **2021**, *8*, 554578. [CrossRef] [PubMed]
9. Roscher, R.; Bohn, B.; Duarte, M.F.; Garcke, J. Explainable machine learning for scientific insights and discoveries. *IEEE Access* **2020**, *8*, 42200–42216. [CrossRef]
10. Samek, W.; Müller, K.R. Towards explainable artificial intelligence. In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*; Samek, W., Montavon, G., Vedaldi, A., Hansen, L.K., Müller, K.R., Eds.; Springer: Cham, Switzerland, 2019; pp. 5–22.
11. Xu, F.; Uszkoreit, H.; Du, Y.; Fan, W.; Zhao, D.; Zhu, J. Explainable AI: A brief survey on history, research areas, approaches and challenges. In Proceedings of the CCF International Conference on Natural Language Processing and Chinese Computing, Dunhuang, China, 9–14 October 2019; Springer: Cham, Switzerland, 2019; pp. 563–574.
12. Lapuschkin, S.; Wäldchen, S.; Binder, A.; Montavon, G.; Samek, W.; Müller, K.R. Unmasking Clever Hans predictors and assessing what machines really learn. *Nat. Commun.* **2019**, *10*, 1096. [CrossRef]
13. Wortham, R.H.; Theodorou, A. Robot transparency, trust and utility. *Connect. Sci.* **2017**, *29*, 242–248. [CrossRef]
14. Rawlins, B.L. Trust and PR Practice. 2007. Available online: <https://www.instituteforpr.org/wp-content/uploads/Rawlins-Trust-formatted-for-IPR-12-10.pdf> (accessed on 9 June 2022).
15. Rawlins, B.R. Measuring the relationship between organizational transparency and employee trust. *Public Relat. J.* **2008**, *2*, 1–21.
16. Liang, L.; Tian, G.; Zhang, X.; Tian, Y. Help comes from understanding: The positive effect of communication visibility on employee helping behavior. *Int. J. Environ. Res. Public Health* **2020**, *17*, 5022. [CrossRef] [PubMed]
17. Scott, S.D.; Carpendale MS, T.; Inkpen, K. Territoriality in collaborative tabletop workspaces. In Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, Chicago, IL, USA, 6–10 November 2004; pp. 294–303.
18. McAllister, D.J. Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations. *Acad. Manag. J.* **1995**, *38*, 24–59. [CrossRef]
19. Fiske, S.T.; Cuddy, A.J.; Glick, P. Universal dimensions of social cognition: Warmth and competence. *Trends Cogn. Sci.* **2007**, *11*, 77–83. [CrossRef] [PubMed]
20. Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An integrative model of organizational trust. *Acad. Manag. Rev.* **1995**, *20*, 709–734. [CrossRef]
21. Christoforakos, L.; Gallucci, A.; Surmava-Große, T.; Ullrich, D.; Diefenbach, S. Can robots earn our trust the same way humans do? A systematic exploration of competence, warmth, and anthropomorphism as determinants of trust development in HRI. *Front. Robot. AI* **2021**, *8*, 640444. [CrossRef]
22. Bellotti, V.; Edwards, K. Intelligibility and accountability: Human considerations in context-aware systems. *Hum. Comput. Interact.* **2001**, *16*, 193–212. [CrossRef]
23. Diefenbach, S.; Butz, A.; Ullrich, D. Intelligence comes from within—Personality as a UI paradigm for smart spaces. *Designs* **2020**, *4*, 18. [CrossRef]
24. Weiser, M. The computer for the 21st century. In *Scientific American*; Springer Nature: Berlin, Germany, 1991; Volume 265.3, pp. 94–105.

25. Ishii, H.; Ullmer, B. Tangible bits: Towards seamless interfaces between people, bits and atoms. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 22–27 March 1997; pp. 234–241.
26. Milgram, P.; Takemura, H.; Utsumi, A.; Kishino, F. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telem manipulator and Telepresence Technologies*; SPIE: Bellingham, WA, USA, 1995; Volume 2351, pp. 282–292.
27. Cao, J.; Lam, K.Y.; Lee, L.H.; Liu, X.; Hui, P.; Su, X. Mobile augmented reality: User interfaces, frameworks, and intelligence. *ACM Comput. Surv. CSUR* **2021**. [[CrossRef](#)]
28. Benko, H.; Wilson, A.D.; Zannier, F. Dyadic projected spatial augmented reality. In Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, Honolulu, HI, USA, 5–8 October 2014; ACM: New York, NY, USA, 2014.
29. Bimber, O.; Raskar, R. *Spatial Augmented Reality: Merging Real and Virtual Worlds*; AK Peters: Natick, MA, USA; CRC Press: Boca Raton, FL, USA, 2005.
30. Fender, A.R.; Benko, H.; Wilson, A. Meetalive: Room-scale omni-directional display system for multi-user content and control sharing. In Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, Brighton, UK, 17–20 October 2017; pp. 106–115.
31. Santos, J.; Rodrigues, J.J.; Casal, J.; Saleem, K.; Denisov, V. Intelligent personal assistants based on internet of things approaches. *IEEE Syst. J.* **2016**, *12*, 1793–1802. [[CrossRef](#)]
32. Hervás, R.; Bravo, J. Towards the ubiquitous visualization: Adaptive user-interfaces based on the Semantic Web. *Interact. Comput.* **2011**, *23*, 40–56. [[CrossRef](#)]
33. Kim, K.; Boelling, L.; Haesler, S.; Bailenson, J.; Bruder, G.; Welch, G.F. Does a digital assistant need a body? The influence of visual embodiment and social behavior on the perception of intelligent virtual agents in AR. In Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Munich, Germany, 16–20 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 105–114.
34. Dhiman, H.; Wächter, C.; Fellmann, M.; Röcker, C. Intelligent Assistants: Conceptual Dimensions, Contextual Model, and Design Trends. *Bus. Inf. Syst. Eng.* **2022**. [[CrossRef](#)]
35. Butz, A. User Interfaces and HCI for Ambient Intelligence and Smart Environments. In *Handbook of Ambient Intelligence and Smart Environments*; Nakashima, H., Aghajan, H., Augusto, J.C., Eds.; Springer: Boston, MA, USA, 2010. [[CrossRef](#)]
36. Sovacool, B.K.; Del Rio, D.D.F. Smart home technologies in Europe: A critical review of concepts, benefits, risks and policies. *Renew. Sustain. Energy Rev.* **2020**, *120*, 109663. [[CrossRef](#)]
37. Jakobi, T.; Stevens, G.; Castelli, N.; Ogonowski, C.; Schaub, F.; Vindice, N.; Randall, D.; Tolmie, P.; Wulf, V. Evolving needs in IoT control and accountability: A longitudinal study on smart home intelligibility. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–28. [[CrossRef](#)]