Exploring the Effects of Head-Mounted Augmented Reality on Helping Behaviour

Jan Ole Rixen
jan.rixen@uni-ulm.de
Institute of Media Informatics, Ulm University
Ulm, Germany

Jan Henry Belz
jan.belz@uni-ulm.de
Institute of Media Informatics, Ulm University
Ulm, Germany

Luca-Maxim Meinhardt
luca.meinhardt@uni-ulm.de
Institute of Media Informatics, Ulm University
Ulm, Germany

Jan Gugenheimer
jan.gugenheimer@uni-ulm.de
jan.gugenheimer@telecom-paris.fr
Technische Universität Darmstadt
Darmstadt, Germany

Enrico Rukzio
enrico.rukzio@uni-ulm.de
Institute of Media Informatics, Ulm University
Ulm, Germany

ABSTRACT

Augmented Reality (AR) can alter environments and steer attention. While prior work dominantly focuses on exploring performances of augmentations, this work aims to understand the societal impact of AR in complex social situations. Focusing on prosocial helping behaviour, we created two scenarios and designed five augmentations aiming to motivate a user to help. We wanted to understand (1) the impact on situation perception and (2) the impact on the social structure. In an online video experiment (n=294), we found that augmenting can impact anxiety about the situation and significantly increase the perceived reason to help being directed towards the augmentation. Similarly, we found that the helped rated the "reason" and "thankfulness" significantly higher towards AR than the helper, creating a disagreement around agency and responsibility. We discuss the implications of AR in complex social structures and how responsibility and agency will become important when embedding AR in our social lives.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in HCI**; **Mixed / augmented reality**.

KEYWORDS

Helping Behaviour, Helping Behavior, AR, Social, Videostudy

ACM Reference Format:
Jan Ole Rixen, Jan Henry Belz, Luca-Maxim Meinhardt, Jan Gugenheimer, and Enrico Rukzio. 2023. Exploring the Effects of Head-Mounted Augmented Reality on Helping Behaviour. In *International Conference on Mobile and Ubiquitous Multimedia (MUM '23), December 03–06, 2023, Vienna, Austria.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3626705.3627969

MUM '23, December 03–06, 2023, Vienna, Austria

ACM ISBN 979-8-4007-0921-0/23/12.
https://doi.org/10.1145/3626705.3627969

1 INTRODUCTION

Until lately, our awareness and assessment of complex social situations could only be attained through our own interpretation of our own sensory stimuli in combination with prior knowledge obtained during our lives. With wearable Augmented Reality (AR) technology, a new dynamic layer of information can now be added to our sensory repertoire. Similar to our sensory perception of the environment, an AR device can observe our surroundings and gather information. Utilizing its connectivity, it can then infer and display information that we might not be able to conduct from our human sensory stimuli and knowledge of the world alone. In supporting impaired users, researchers have already used artificial intelligence in combination with AR technology to support users by helping them interpret their surroundings through head-mounted AR devices [15, 56].

Wearable AR devices of the current and especially the following generations will most likely not be used in isolation but embedded into their user's daily life [31, 68, 71, 72]. While prior work has already established the concept of augmenting the counterpart in a conversation with previously recorded personal information about the actors [1, 28, 44], others argue that it could also use artificial intelligence to infer new information from those situations [7, 71]. Building on this, we argue that like helping impaired users to interpret a facial expression [15, 56], future wearable AR Devices and their ability to interpret could also be used to make users aware of certain situations happening in their surroundings. They could interpret such information for them and might even be used to guide the user through their decision process of interacting with this situation. This, in turn, raises the question of how such interventions could look, how effective they can be, and how they would impact our social fabric.

To explore this, we look at one highly discussed kind of prosocial behaviour impacting our social fabric: helping behaviour in daily life and the so-called "bystander effect." Intervening in potentially dangerous situations a vast amount of what is called moral courage [30, 38] and, therefore, the willingness to betake oneself to risk one's health. While some situations need a person to step up, general help-giving can also include prosocial acts that do not necessarily require extensive courage [38], e.g., opening or holding open a door for others. In any helping behaviour case, bystanders need to go through a decision process on whether to intervene, impacted by

numerous factors. Deciding for or against helping behaviour, in this regard, combines considerations about the situation (e.g., its ambiguity [13]), its assessment (feeling of responsibility [29, 47, 65] and anxiety [21]), and factors individual to the person and their current situation (e.g., mood [38]).

With our work, we want to give first insights on how the coming AR technology [42] could influence their users' perception of and feelings towards a helping situation, as well as get first insights into how it might influence how different actors evaluate such situations. We investigate this in the context of helping behaviour in a help-giving situation (a person not getting through a door) and a situation needing moral courage (a person being attacked). In this context, we conducted a mixed factorial online video experiment (n=294) to investigate augmentations based on the five levels of intervening in a potential helpers decision process as defined by Latané and Darley [50]: (1) creating Attention, (2) creating Emergency Awareness, (3) Attributing Responsibility, (4) showing the necessary skills to help, and (5) making the final decision for them. We investigate how augmentations influence how both helper and person in need feel about the situation and potential help in both moral courage and help-giving situations. By investigating the impact on the perception of the situation, we want to receive first hints on the effectiveness of such intervention.

Our results show that augmenting a situation for awareness made participants less anxious and worried compared to not augmenting it all. In turn, further augmenting it with, e.g. instructions to help vary in its effect on the participants. While such augmentations led to lower levels of anxiety in the non-threatening help-giving situation, they fostered anxiousness in the threatening moral courage situation. Regarding the social impact of augmentations in helping behaviour situations, our results suggest that AR devices can be perceived as influential actors in such situations. How much influence is attested herein varies with the role, with the helped person attesting a greater influence than the helper. We discuss this difference in influence attribution and how it can influence social structures when AR devices are perceived as able to support prosocial behaviour like helping. We also discuss the general ethical implications of such augmentations.

2 RELATED WORK

In our work, we build on multiple fields of research. First, Augmented Reality in a social context as well as help-giving, and the decision progress in deciding for or against help-giving. In the following, we will introduce these topics by laying out previous work.

2.1 Social AR

Prior work in the field of AR in social contexts has shown that users can be supported by Augmented Reality technology in multiple ways. One way of those ways is to foster interactions between individuals by displaying additional information in the user's field of view. Firstly, researchers tried to make introductions easier by showing additional information that might make people more approachable. McCarthy et al. [59] used RFID chips to identify bypassing persons at a scientific conference, displaying their name, affiliation, and picture on a public screen to kick-start conversation. Other researchers asked participants to create digital self-representations [45, 46, 60], which then were displayed to ease getting into a conversation. Additional information could, e.g., be derived by recognizing the user's face and identifying them in the user's contact list [44], or on social media [1]. Like starting a conversation, such technology could also be utilized to keep it going. This can either be done by displaying interests both interlocutors share [62, 79] or by creating further interest with topics that are not mutual [45].

In contrast to querying existing data, Bermejo et al. [7] argue that Augmented Reality and Artificial intelligence in the form of Big Data "have a logical maturity that inevitably will converge" [7, p.1]. Adding to this argument, Rixen et al. [71] argue that current and future advancements in AI technology could also be utilized to analyze the user's current context by using the device's sensory equipment . With AI starting to outperform humans in a multitude of areas (e.g., detecting emotional state through micro expressions[54], person's age [19], sexual orientation or race [26] through visual clues [78]), the combination of AR and AI could reveal new information about the user's situation to them. Previous work has also started exploring how AR technology can aid people by interpreting their surroundings in a context for them. Daniels et al. [15] and Liu et al. [56] analyze the current situation and aid children with an autism spectrum disorder in behaving by textually displaying their social counterpart's current emotional state.

2.2 Prosocial Behaviour: Help-Giving, Moral Courage and Heroism

Prosocial behaviour is an umbrella term used for "positive social acts carried out to produce and maintain the well-being and integrity of others" [9, p.710]. Such behaviour can include helping, sharing, donating, cooperating, and volunteering. Each prosocial behaviour thereby involves interactions between at least two actors. A Benefactor (or helper) and a person that is being helped [38].

Three types of behaviour that fall under this term of prosocial behaviour and are classified as helping behaviour are help-giving, moral courage, and heroism. Moral courage is defined as a "morally brave and risky behavio[u]r [...] without weighing potential disadvantages" [38]. Individuals that are showing moral courage act to their moral standards while disregarding the risk of negative consequences. Kinnunen and Windmann [40] define moral courage as a form of altruism as the behaviour is costly to oneself while benefiting others. This only partly applies to the concept of heroism as described by Osswald et al. [65]: While a helper in a moral courage situation faces negative social consequences through the perpetrator, a hero does not act entirely selflessly but also expects positive social consequences, e.g., applause and admiration by another person. Help-giving behaviour, in turn, more generally describes the act of assistance, which does not necessarily require courage (e. g. holding a door) [38]. Osswald et al. [65] distinguish between moral courage, help-giving, and heroism in terms of the social consequences a potential helper has to face: while help-giving usually is rewarded with social acknowledgement without requiring the risk of negative social consequences, moral courage stands for intervening despite facing such potential negative social consequences [22, 29]. Heroism, in turn, incorporates both negative and positive social consequences. They,e.g. face the direct negative consequences of intervening in a dangerous situation but also expect positive consequences by being praised for it.

2.3 Decision to Help

If one gets into a situation where one can decide for or against helping, multiple mental steps must be taken to arrive at a final decision. A theoretical process model that describes this process is the five-step process model of helping introduced by Latané and Darley [50]. In the following, we will discuss these five stages (Attention, Emergency Awareness, Attribution of own Responsibility, Skills for Helping, and Final Decision; please see Figure 1 for a visual depiction of the model) and discuss the psychological processes that can interfere and prevent reaching the final decision to help. In the context of the bystander effect, Latané and Darley [50] define three main psychological processes, namely Pluralistic Ignorance, Diffusion of Responsibility, and Evaluation Apprehension. We will also discuss these while laying out the process model of helping. Apart from this model and the bystander effect, we will also discuss general influences on helping behaviour.

To evaluate an incident and come to a sentiment and decision to act, a potential helper must first become aware of it. This gaining of situational attention is the model's first step, which we will call *Attention*. Having gained attention, a potential helper proceeds by assessing the situation and constituting it as an emergency (*Emergency Awareness*). At this second step, Rendsvig [69] locates the first of three major psychological processes interfering with potential helping behaviour, which Latané and Darley [50] have identified as pluralistic ignorance. *Pluralistic Ignorance* happens in an ambiguous situation, where the potential helpers "may choose to seek social proof in order to individually determine a correct course of action and the associated consequences thereof" [69, p.2]. Unsure about the situation's urgency and relying on the interpretation of non-helping others, a potential helper can conclude that no one else perceives an emergency and, therefore, non exists [51]. A high amount of ambiguity furthers this occurrence [8, 13]. Latané and Rodin [52] found that ambiguity of emotions displayed by other bystanders furthers the reluctance to help, as helping behaviour was more common when being with known persons that are easier to read compared to strangers. In their study, Solomon et al. [75] reported that relying on others' interpretation also could work towards the opposite outcome, as only one person helping made others acknowledge an emergency and led to more people helping [75].

In the following third step, *Attribution of Responsibility*, the potential helper must develop and accept a feeling of responsibility. Here, the second major interference can occur, which Latané and Darley [50] identify as *Diffusion of Responsibility*. Having one or more other bystanders, the individual will only feel responsible for a part of the potential happenings in the case of non-intervention [23]. Making it less probable for the individual to intervene. In this context, Latané and Rodin [52] found that groups of two were less likely to help an injured woman than when being alone. In another study, Latané and Dabbs Jr [49] found that this inhibition to help increased with the number of people present and is influenced by a bystander-group's characteristics [73] and the individual interpretation of other bystanders' ability to help [8].

In the fourth stage, Skills for Helping, the potential helper evaluates if they possess the skills necessary to succeed in helping. Here, the last of the major interference may happen, which Latané and Darley [50] identify as *Evaluation Apprehension*. It refers to the fear that

Figure 1: The proposed process model of helping by Latané and Darley [50].

others might judge one who acts publicly. Here, they fear making a mistake or acting inadequately while being observed, hindering their decision to help [24]. Having also traversed through this step, a potential helper arrives at the fifth and last step. They must finally reach a conscious decision to help and act on it.

While *Pluralistic Ignorance*, *Diffusion of Responsibility*, and *Evaluation Apprehension* influence a person mainly through the presence of other bystanders, other determinants can be attributed to victim-related factors or internal factors of the potential helper. One of the main factors that have to be overcome is the anxiety felt when encountering an incident in which help can be given [29, 47, 65]. It was also shown that people possessing certain attributes are more likely to help. Witnesses that perceive themselves as stronger, more aggressive [37], or more empathetic Laner et al. [48] are more likely to intervene and help. Also, physical attributes like height [48] or a person's biological sex [49] were shown to influence the decision. Other influencing factors are the attributes of the person in need themselves. Here, e.g., their perceived biological sex [49] and their relationship towards the potential helper [66] influences the final helping decision.

3 RESEARCH QUESTIONS

As Augmented Reality and artificial intelligence are destined to converge bermejo2017augmented, and researchers start to include AR in a social context [e.g. 1, 28, 44] and help users to interpret their current situation [e.g. 15, 56], this work asks the question of what would happen if complex social situations are interpreted and augmented for their users. We, therefore, use helping behaviour as an example for general prosocial behaviour and try to get first insights into (1) how augmenting such situations can influence a person's perception of a potential helping situation. We also want to explore how (2) the helper and person in need perceive the helping behaviour and the AR device's influence on it. In the following, we will lay out our research questions and describe how augmentations could influence situational perception and feelings linked to a person's helping decision. Laying out or reasoning, we will also present possible augmentations intervening in each step in the five-step process model of helping introduced by Latané and Darley [50]. We will begin with our first research question regarding a person's perception of a potential helping situation.

> ***RQ1:*** *Can AR systems in potential helping situations influence the (1) Situation Assessment and (2) Feelings induced by the situation?*

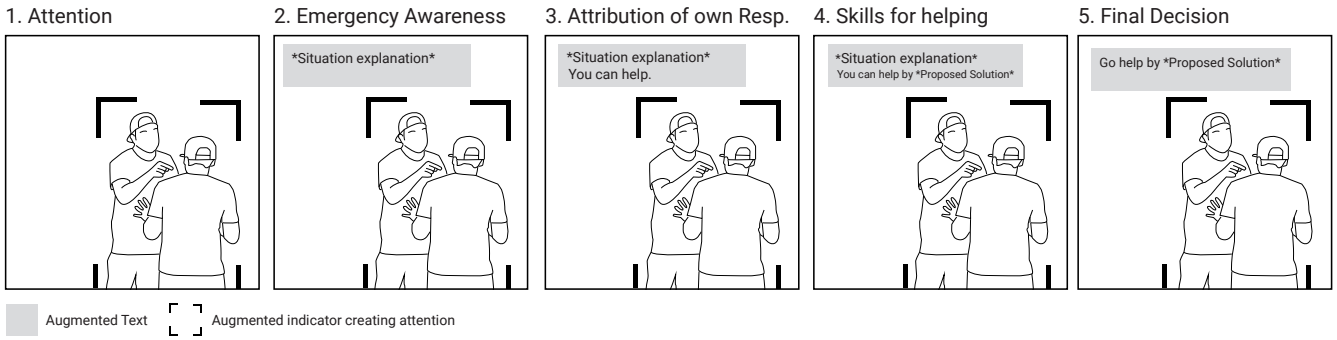


Figure 2: Augmented interventions for helping behaviour situations as described in Figure 4.2. Each intervention corresponds to one level of the process model of helping proposed by Latané and Darley [50].

First, we were interested in whether and how augmentations can influence the intervening factors that arise from the bystander effect. *Pluralistic Ignorance*, as described in subsection 2.3, originates in the ambiguity of the situation and the missing proof of others expressing that they also perceive the situation as indeed a situation in which help is appropriate. With AR devices, having an entity that identifies and defines the situation could help combat this ambiguity. Here, the AR device could act like an additional bystander expressing the acknowledgment of the emergency, which could motivate the potential helper to act. This could be done by not only highlighting the situation (Step 1) but also describing it in a way that implies the emergency (Step 2). We will call this description "*Situation Explanation* " (see Figure 2). Further, intervening on a deeper level, openly stating the potential helper's responsibility to help (Step 3), could, in turn, help combat the diffusion of responsibility. *Evaluation Apprehension* arises from the feeling of not knowing how to adequately intervene in an incident without being negatively judged about it later. Here, the AR device could help in two ways. First, it could propose an adequate solution (*Proposed Solution*) so the potential helper knows how to react without being judged. Second, the potential helper could know they have a target to blame if the help is perceived negatively in hindsight. As anxiety was a re-occurring hindering factor in related work about helping behaviour, we were also interested in whether the potential helper's anxiety could be influenced by not being alone in the situation but supported by the device (in Steps 1 and 2) or even negatively influenced when the own responsibility is exposed, or the final decision to help is presented. Therefore, the (1) Situation Assessment part of the research question refers to assessing if augmentations can influence *Pluralistic Ignorance*, *Evaluation Apprehension*, and the sense of responsibility. The second part, (2) Feelings, in turn, focuses on the participant's sense of anxiety in the situation. Secondly, we were interested in how both the helper and the person in need perceived the AR device's influence on actual helping behaviour. Here, we were interested in whether helpers still felt full agency over their helping behaviour or would attribute parts of their decision process to the AR device and its augmentations. In this context, we suspect that the helper and person in need could have differing opinions on how the AR device and its augmentations influenced behaviour and which role it played in the final decision to help. In addition to comparing both parties' assessments of how much agency the AR device takes from the helper, we also set out to compare assessments on how grateful the person in need is towards the Augmented Reality device and if the *Helper* would have helped without the AR device. This leads us to our second research question:

> ***RQ2:*** *How does augmentation influence the perception of helping behaviour and the device's impact on it?*

As Rixen et al. [71] argue, to make AR a space where every actor feels comfortable and socially accepted, it has to be ensured that anyone involved feels comfortable with the augmentations. Therefore, we want to explore how helpers and persons in need feel about the augmentations and to reveal in which situation they feel comfortable with it.

4 METHODOLOGY

In the context of helping behaviour, we utilize videos to get first insights into how augmenting prosocial situations can influence a person's willingness to help, their perception of the situation, and the perceived influence of the AR devices from the perspective of helper and person in need. Inspired by and in line with prior work by, e.g., Ma et al. [57] and Rixen et al. [71], we used a mixed factorial design, including two between-subject factors as well as one within-subject factor (*Situation*). With this approach, we arrived at twelve between-subject conditions built upon the between-subject factors *Role* and *Intervention*. In each condition, we showed mock-up videos to the participants, asking them to imagine themselves in the *Role* of either one part of every prosocial situation: the potential *Helper* or the helped [38]. While being in the context of helping behaviour, we define this person as "person in need " (*PiN*). With *Situations* being the within-subject factor, participants were confronted with both *Help-Giving* and *Moral Courage* situations. The situations were augmented depending on a participant's condition, building on the decision model introduced by Latané and Darley [50]. Leading to the five levels of *Intervention*: *Attention*, *Awareness*, *Responsibility*, *Skills*, *Final Decision*, and a condition in which the situation was not augmented (*None*). This means that each participant imagined themselves in one of the *Roles* and was then exposed to both *Situations* augmented with the same specific *Intervention*. Following, we will present the video artifact used for our study and lay out the rationale behind the choices that led to the final result.

4.1 Reasoning Behind Conducting an Online Experiment

In accordance with prior studies [43, 71, 72], we avoided an artificial lab study by choosing the approach of showing participants specific situations and letting them imagine themselves being in them. By doing this, we could rule out accompanying biases that the study would have suffered otherwise. For example, conducting a study with today's clunky AR headsets that can only display in a restricted part of the user's field of view would lead to hardware biases, making the futuristic-looking functionality less believable. Using a video-based approach, we could create a credible scenario and expose each participant to the exact same scenario, which would be challenging to recreate similarly and believably when executed in a real-life situation.

A similar type of imaginative study design is used in social science research. With experimental vignette studies (EVS), researchers are "presenting participants with carefully constructed and realistic scenarios to assess dependent variables including intentions, attitudes, and behaviours" [p. 352 3]. Such a scenario, or "vignette", is "a short, carefully constructed description of a person, object, or situation, representing a systematic combination of characteristics" [p. 128 5]. The vignettes are not limited to consisting of textual information but can also include the usage of videos and other types of media [36]. They also enable participants to participate in their own context without needing to travel to a laboratory [74], which, in turn, enables them to reach a larger audience (in this case, N=294). According to a review by Aguinis and Bradley [3] this type of research design has been in existence for several decades and has been widely applied in fields such as business ethics. Here, we want to highlight that we are not exploring actual behaviour in our depicted scenarios, as this would be difficult due to the so-called Intention–Behaviour Gap [10, 33]. Instead, we want to research the impact on situation perception and the impact on the social structure. In the context of helping behaviour, video footage has also been utilized before [53]. Analogous to our study, participants subsequently completed a questionnaire about the assessment of the incident, their emotional response towards it, and their likelihood to intervene if they faced such a situation. Building on this prior work and established approach, we decided to confront participants with the augmented *Help-Giving* and *Moral Courage* situation in the form of mock-up videos we created using Adobe Aftereffects and Mocha AE for planar tracking.

4.2 Apparatus

To allow participants to imagine themselves in this situation and perceive our augmentations, we shot the videos from the point of view of a person wearing a future AR device. This person is on their way home from an undefined task. On this journey through a building, they are assisted by the features of their AR glasses. After an approximately 50-second prelude to introduce and immerse them in a future world with AR augmentations, they encounter either a *Help-Giving* or *Moral Courage* situation that is augmented as described in section 3 and is depicted in Figure 2. In the following, we will lay out the reasoning behind the choices of *Help-Giving* and *Moral Courage* situation, the presence of a bystander, and the general non-incident-related augmentations. We will start with the choices of scenarios for a *Help-Giving* and *Moral Courage* situation that both could be perceived for approximately 15 sec before the video ended while still being in the situation.

For the *Help-Giving* situation, we had to create a scenario solvable without getting the potential helper into danger or yielding negative social consequences for them. Peter et al. [67] investigated the bystander effect in such a non-threatening, non-emergency situation by having a person knock on a door that a participant willing to help could open. We chose to adopt this method. In our context, instead of having auditory cues that one can help by opening a door, we showed the participants visual ones. This reasoning resulted in a scenario with a person carrying a tablet stacked with coffee, cakes, and the corresponding tableware. A screenshot of how this situation looked in the final video can be seen in sub-figure c of Figure 3. As *Situation Explanation*, we used "There is a person not getting through the door." which we argue describes the happening situation fittingly. As *Proposed Solution*, we augmented that the viewer might help them "by opening the door for them". Asking the participants of our study who perceived no augmentation how they would help the person in need, all responded that they would have opened the door for the person, which again confirms our choice of *Proposed Solution*.

In contrast to a *Help-Giving* situation, a *Moral Courage* situation has to include some danger and the implication of possible negative social consequences for the potential helper like "being insulted, excluded, attacked, psychically or physically violated" [65, p.393] (see subsection 2.2). Cases of moral courage (e.g., a female student saving Syrian refugees from aggressors) [1] often include one or more physically stronger aggressors physically threaten or harm their victims until a helper intervenes. Going with this archetype, we decided to display a situation depicting a person getting stopped and physically threatened and harassed by a bigger aggressor. To visually support the situation, we dressed the aggressor in black, a colour associated with strength [2], and aggressiveness [25]. A screenshot of the situation can be seen in sub-figure b of Figure 3. As *Situation Explanation*, we decided to use "There is a person getting attacked by an aggressive person.". Since we argue that a potential future system would try to avoid putting its users in danger and adhere to official guidelines, we based the *Proposed Solution* on such official guidelines. The German police [2] published a guide on how to show moral courage and intervene in critical situations without getting into danger. Here, they suggest speaking directly to other (possible) helpers or loudly declaring to organize help, which might already lead to the offender letting go of their victim. We, therefore, arrived at the *Proposed Solution* of helping the victim "by loudly requesting the bystander to help the person in need together".

As described above, major factors meddling with helping behaviour are the three bystander-related psychological processes defined by Latané and Darley [50]. As we wanted to explore how augmentation influences those psychological processes, a bystander had to be present in our video. We also decided to have the bystander act passively and uninterested while they visibly look at the situation. To make it easier for participants to imagine them-self in the place

[1] https://www.freiepresse.de/chemnitz/studentin-rettet-syrer-vor-rassistischen-schlaegern-artikel9448997, ACCESSED: 08-DECEMBER-2022
[2] https://www.bundespolizei.de/Web/DE/02Sicher-im-Alltag/02Zivilcourage-zeigen/Zivilcourage-zeigen_node.html, ACCESSED: 30-AUGUST-2022

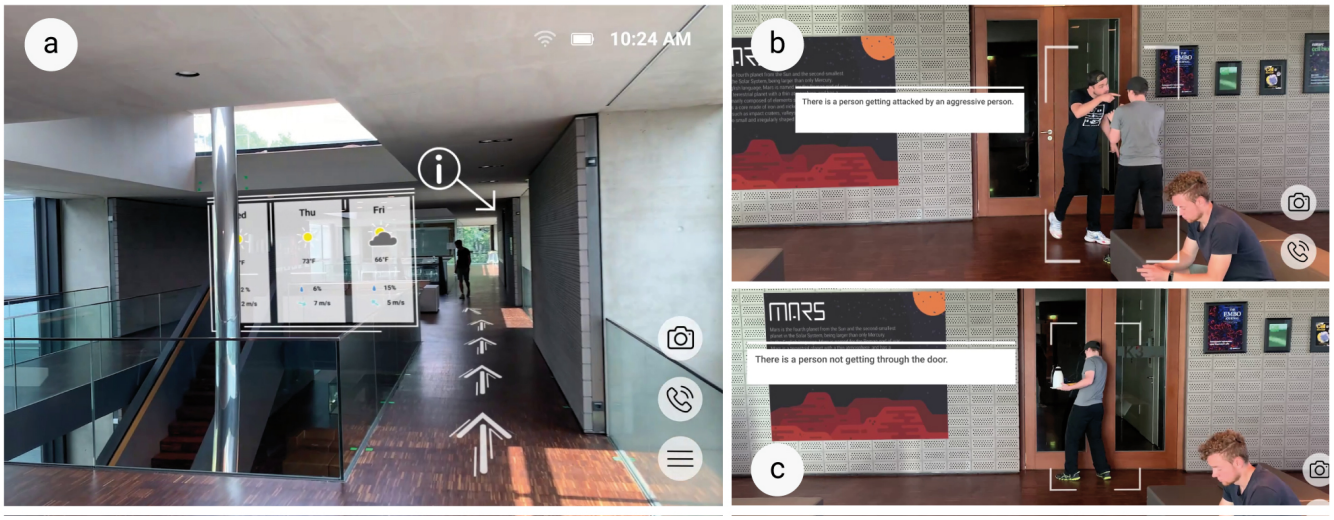


Figure 3: Screenshots from videos shown to the participants of our experiment. The left picture (a) shows a screenshot from the intro section that proceeds either the augmented *Moral Courage* (b) or *Help-Giving* situation (c). This screenshot (a) shows how the UI, navigation, and weather forecast were anchored in their surroundings. The screenshots on the right depict how augmented interventions looked for *Moral Courage* situation (b) and *Help-Giving* situation (c) with the *Intervention* of *Awareness*, which corresponds to the second step of the process model of helping proposed by Latané and Darley [50].

of the person in need, we chose an actor with an average stature not perceived as unusually tall or trained. To build a more dense AR scenario and not have the incident-related augmentation stand out as the only one, we populated the video with further augmentations depicted in a prelude embedded in the overarching story: The person is on their way home supported by their guiding navigation app [4, 34, 61]. On their way towards the incident, the viewer encounters an artwork [11] and trees [20] augmented with further information as well as learning-focused information [39] about the planet Mars. To create a denser surrounding with augmentations, we added further augmentations displaying the weather forecast anchored in the surroundings and an incoming call. The full videos can be perceived in the supporting material added to this work.

4.3 Procedure

We structured the study into four sections. After study registration and accepting the consent form, we presented participants with those sections in the following order.

Part 1: Individuality. From prior work, we already know that individual traits can influence how a person reacts to a helping situation. Therefore, we first queried participants' individual features that have been shown to impact helping behaviour. Following prior work by Huston et al. [37] and Laner et al. [48], we asked participants to state if they would agree with being *stronger*, *more aggressive*, and *more sympathetic* than other people. In line with the mentioned prior work, we also measured it on a scale from 1 (strongly disagree) to 5 (strongly agree).

Part 2: Introduction to AR. Following the first questions, to ensure a general understanding of Augmented Reality technology and its ability to analyze a user's surroundings, we gave our participants an introduction to those topics. In line with prior work by Rixen et al. [71, 72], we first introduced the participants to the concept of *AR-HMDs* by providing textual information supported by a mock-up video. After showing them how such a device could look and display information, we further explained its ability to analyze a user's surroundings and display the results to them. Here we focused on making the participants understand that this does not only include information about objects but the actions and feelings of other persons.

Part 3: Primary Video Conditions. After participants were familiarized with Augmented Reality technology, the main part of the study began, confronting participants with both a *Moral Courage* and a *Help-Giving* situation in the form of a mock-up video (for more details, see Figure 4.2) which included augmentations according to the *Intervention* of their condition. Before seeing the first video, we introduced the participants to the general situation depicted in the videos. They then got introduced to their task of imagining being the POV (*Helper* condition) or a person dressed in the light grey shirt (*PiN* condition). To make it easier for participants of the *PiN* condition to identify the person they have to imagine being, we added screenshots of the person to the explanation. To ensure the participants understood their task, we asked comprehension questions and excluded those failing them from the evaluation. After having this assurance, participants were presented with *Moral Courage* and *Help-Giving* videos in a randomized order to avoid carryover effects. The questionnaires belonging to a condition were further separated into two groups. Each video ended after the participant had been exposed to the augmentation for 11 seconds, still looking at the

Who was Asked	Asked when...	Variable Measured	Measuring influence on...
Helper	...in the situation	*need to help*	... Situation Assessment (**RQ1**)
		responsibility to help	
		skills for helping	
		willingness to help	
		anxiety worrying nervousness	... Feelings (**RQ1**)
	... situation has been resolved	*agency**	
		*influence on helping decision**	
		*main reason to help**	
Helper and *PiN*	...in the situation	*comfort*	... Feelings of Comfort (**RQ2**)
	... situation has been resolved	gratitude	... Perceived Influence of Technology (**RQ2**)
		help without device	
		agency*	
		influence on helping decision*	
		*main reason to help**	

* asked for Helper and PiN but also evaluated for Helper alone

Figure 4: The measured variables for only *Helper* and *Helper* and *PiN*. It is indicated if the variables were measured while imagining being in the situation or while imagining that the situation was resolved. Also the related RQ is indicated.

situation without it being resolved. In the first set of questions, participants had to imagine still being in this situation they just have seen. Afterward, participants were told to imagine that the *Helper* helped and resolved the situation, supported by the augmentations. This was followed by the second set of questions. Which variables we measured and how we measured them will be described in subsection 4.4.

Part 4: Demography. Ending the study, we queried demographic data. As Laner et al. [48] also found significant differences in the helper's height, we made sure also to query it.

4.4 Measurements

In the following, we will describe our dependent variables and how we measured them. To make those measurements easier to understand we arranged them depending on who was asked (only *Helper* or *Helper* and *PiN*) and if participants were still imagining being in the situation or that the situation had been resolved. We also relate the variables to our research questions. A visual overview can be seen in Figure 4.

***Helper*.** We will first cover the measurements that belong to RQ1 which deal with the influence of the augmentations on the *Helper*'s Situation Assessment and Feelings. While participants in the *Helper* condition were still imagining being in the situation, we measured multiple dependent variables for them only. First, those responding to the Situation Assessment (RQ1) and, therefore, especially the bystander effect (see subsection 2.3). Pluralistic Ignorance shows through a potential helper being confused about whether person in need needs help or not. We, therefore, asked participants to state their agreement with the statement "The person dressed in the light grey shirt needs help" (*need to help*). This item and the following were rated on a Likert Scale from 1 (Strongly Disagree) to 7 (Strongly Agree). Diffusion of Responsibility shows through not feeling responsible for intervening in a helping behaviour situation when others could also help. We took an item that Fischer et al. [21] used to determine the feeling of social responsibility. Adopting it to our situations, we asked participants to state their agreement towards "I feel personally responsible for helping the person in the light grey shirt" (*responsibility to help*). Evaluation Apprehension originates in fear of making a mistake or acting inadequately, subsequently leading to the judgment of observers [24]. We, therefore, asked participants to state their agreement towards "I know how to help the person in the light grey shirt in an appropriate way so that I am not judged negatively later" (*skills to help*). We also asked the participants how willing they were to help the person in the observed situation by stating their agreement towards "I would help the person in the light grey shirt" (*willingness to help*).

Responding to the Feelings induced by the situation (RQ1) we measured the feeling of anxiety. Here, we used questions analogous to Baker et al. [6]'s Anxiety Symptoms Questionnaire and asked for the anxiety symptoms. Analogous to their questionnaire, we let participants rate how intense or bothersome the symptom(s) have been imagining themselves in the seen situation on a scale from 1 (None) to 10 (Extreme distress). We also decided to query *worrying*, and *nervousness* from the same questionnaire.

After imagining the situation has been resolved, we measured additional metrics. We also asked *Helper* about their feeling of agency and, respectively, participants in the *PiN* condition about their perception of how much agency the *Helper* has over the situation. For this measurement, we took questions from the Sense of Agency Scale [77] that fitted our situation the most. The first question related to the sense of agency (*SoA*) the *Helper* had, while the second asked about the reversed concept, the sense of negative agency (*SoNA*). For *SoA* we, therefore, asked the participants to state their agreement to "I am in full control of what I do" on a Likert-Scale from 1 (Strongly Disagree) to 7 (Strongly Agree). Similarly, for *SoNA* we asked participants to state their agreement to "I am an instrument in the hands of somebody else". For the evaluation, we calculated an agency score (*agency*) containing both *SoA* and *SoNA* by calculating the mean of *SoA* and the inverted value of *SoNA*. We also asked the *Helper* to state on a Likert-Scale from 1 (Strongly Disagree) to 7 (Strongly Agree) how much they agree with the AR device influencing their decision to help (*infl. on help. decision*: "The AR device influenced my decision to help") and being the main reason why they helped (*main reason*: "The main reason I helped was the AR device").

***Helper* and *PiN*.** We will now cover the measurements that belong to RQ2 which deals with the differences between the assessment of *Helper* and *PiN*. Therefore, all values were measured for *Helper* and *PiN*. While participants still imagined being in the situation, we measured *comfort* for both *Roles*. Analogous to previous studies [12, 57, 71, 72] we measured *comfort* on a Likert-Scale from 1 (Strongly Disagree) to 7 (Strongly Agree) to the question "I feel comfortable with the augmentation involving me (/the person in the light grey shirt)".

After imagining the situation has been resolved, we queried how much *gratitude* the *PiN* had towards the AR device and how the *Helper* estimated it by measuring agreement (Likert-Scale from 1 (Strongly Disagree) to 7 (Strongly Agree)) towards the questions "I'm grateful to the AR device" (*PiN*) and "The person in the light grey shirt is grateful to the AR device" (*Helper*). We also asked if *PiN* and *Helper* thought that the *Helper* still would have helped without the AR device (*would have helped*). To explore differences in perception, we also opposed values that we already talked about in prior. In the same way as above, we, therefore, measured *agency*, *infl. on help. decision*, and *main reason* for the *PiN*, each relating to the *Helper*. So we e.g., let *PiN*

rate "The person wearing the AR device is in full control of what I they do" for *SoA* instead of "I am in full control of what I do" for the *Helper*.

4.5 Participants

For this online experiment, we recruited participants through the Prolific platform[3]. To avoid confounding variables such as culture [70], we recruited US citizens only. We paid our participants an hourly wage of £9 to compensate for their efforts, resulting in a reward of £1.95 for ≈ 13 min in the *PiN* conditions and £2.25 for ≈ 15 min in the *Helper*. We executed the study following the local ethical requirements of the hosting institution.

Initially, we received 322 responses. Of those participants, we had to exclude one for failing the attention checks (designed true to the Prolifics guidelines on fair attention checks [4]) and 28 for failing our comprehension test. Reaching a total of 293 participants. Even though they were excluded, all participants were rewarded for their efforts. This process leads to 24 to 25 participants per condition with an overall age between 18 and 75 (M=36.09 years, SD=13.19). Of those participants, 146 identified as female, 141 as male, three as non-binary, and three preferred not to say.

As described in subsection 2.3, related work has established that various personal characteristics can influence the willingness to help. Due to our between-subject design, group variations defined by their conditions could vary in those characteristics and influence our findings. To rule out such influences as well as possible, we computed Bayes factors for these characteristics. The Bayes factors were computed with the 'BayesFactor' package in R and showed no evidence for the role model neither for perceived own aggression (BF = 0.066), sympathy (BF = 0.18), and strength (BF = 0.45) as well as the participants' height (BF = 0.068). We, therefore, conclude that we found no evidence for heterogeneity and consider them similar in those influencing factors.

5 RESULTS

In the following, we report the results of our experiment ordered by the research questions defined in section 3. First, the measurements regarding *RQ1*, only regarding the influences on the *Helper*. Then, we follow with the measurements regarding *RQ2*, concerning both *Helper* and *PiN*. In this context, we will look at the results of *comfort* separately at the end of this section.

For all following tests, we run Linear mixed models (LMMs). We fitted the LMMs using the lme4 package in R (estimated via restricted maximum likelihood (REML) and nloptwrap optimizer). P-values were computed using a Wald t-distribution approximation. The main benefit is the possibility of specifying random effects. Even though we found no evidence for heterogeneity within the participants' groups on variables specified by related work, we still suspect that the answers might be highly personal. Specifying the participants as random effect allows us to define the portion of variance that is related to these personal differences (example formula: dependent variable *Role* * *Situation* + (1|ParticipantID)). On the basis of empirical data of previous studies Norman [63] argues that parametric tests are robust enough to be used with Likert-type data. LMMs are frequently used and recommended to analyze Likert-type data [e.g. 14, 27]. $R^2conditional$ and $R^2marginal$ were calculated using the report package (Version 0.5.5) in R.

5.1 *RQ1*: Can AR systems in potential helping situations influence the (1) Situation Assessment and (2) Feelings induced by the situation?

As measurements regarding *RQ1* only included participants in the *Helper* conditions, the following evaluation relates to the analysis of 147 participants. The first part will present the results of *Intervention*, and *Situation* influenced the *Helper*'s (1) Situation Assessment and (2) Feelings induced by the situation.

(1) Situation Assessment. We fitted LMMs to predict *responsibility to help*, *need to help*, and *skills to help* with *Situation* and *Intervention*. To account for deviating tendencies in how to interpret and react to an incident, our models included the participants as a random effect. Running those models, they showed no significant main effect of *Intervention*, nor did they show statistically significant interaction effects regarding *Situation* and *Intervention*. Meaning that we found no evidence supporting the influence of *Intervention* on *responsibility to help*, *need to help*, and *skills to help*. Nevertheless the models for *need to help* (R^2 = 0.42, marginal R^2 = 0.04, Intercept at 6.37 [t(280) = 29.03, p < .001]) and *skills to help* (R^2 = 0.51, marginal R^2 = 0.40, Intercept at 6.17 [t(280) = 21.72, p < .001]) showed statistically significant negative main effects of the *Situation* [Ref:*Help-Giving*] on *need to help* (beta = -0.50, t(280) = -2.07, p < 0.05) and *skills to help* (beta = -2.17, t(280) = -5.95, p < .001). This means that participants showed a statistically significant lower agreement to the *PiN* needing help in the *Moral Courage* Situation (M=5.96,SD=1.15) compared to the *Help-Giving* Situation (M=6.36,SD=0.97) and also agreed less to know what to do without being judged in the *Moral Courage* Situation (M=4.13,SD=1.64) compared to the *Help-Giving* Situation (M=6.33,SD=1.11).
Another LMM (R^2 = 0.41, marginal R^2 = 0.19, Intercept at 5.63 [t(280) = 23.12, p < .001]) predicting reported *willingness to help* with *Situation* and *Intervention* showed a statistically significant negative main effect of *Situation* (Ref: *Help-Giving*) on *willingness to help* (beta = -1.33, t(280) = -4.53, p < .001) but no other significant effects. This means that participants expressed a statistically significant higher willingness to help in the *Help-Giving* Situation (M=5.38, SD=0.98) than the *Moral Courage* (M=4.27, SD=1.36) Situation.

(2) Feelings. In the same way, as for (1) Situation Assessment above, we fitted another set of LMMs to predict *anxiety*, *worrying*, and *nervousness* with *Situation* and *Intervention*. The results of the models for *anxiety* and *nervousness* can be observed in Table 1 (we excluded the *worrying* model as it did not yield any significant effects). While we found no significant main effects for the other *Interventions*, we found a statistically significant negative main effect of *Attention* (Ref: *None*) on *anxiety* and *worrying*. This means that participants reported statistically significant lower levels of *anxiety* when the *Situation* was augmented for *Attention* (M=3.69,SD=2.69) compared to having no *Intervention* (M=4.77,SD=2.67). This, also, means that

[3]https://www.prolific.co/, Accessed: 10-August-2022
[4]https://researcher-help.prolific.co/hc/en-gb/articles/360009223553-Using-attentionchecks-as-a-measure-of-data-quality, Accessed: 10-AUGUST-2022

participants reported being less worried when the *Situations* were augmented for *Attention* (*M*=3.41,*SD*=2.7) compared to having no *Intervention* (*M*=4.79,*SD*=2.58).

We found a statistically significant and positive main effect of *Situation* (Ref: *Help-Giving*) on *anxiety* and *worrying*. This means that participants reported higher levels of anxiety and worry when observing the *Moral Courage* compared to the *Help-Giving* Situation. Additionally, we found that the *Interventions Awareness*, *Responsibility* and *Skills* (all having *None* as reference) showed statistically significant positive interaction effects with *Situation* (Ref: *Help-Giving*) regarding the reported level of *anxiety*. This means that augmenting the *Moral Courage* Situation with a *Awareness*, *Responsibility* and *Skills Intervention* led to higher levels of *anxiety* compared to the *Help-Giving* Situation. Please refer to Figure 5 for a visual depiction as well as M and SD values. Looking at the visual depiction, a cross-over effect can be observed. While *Intervention* in the *Help-Giving* Situation led to lower *anxiety* levels than those reported when no augmentation (black line) was present, in the *Moral Courage* Situation, *anxiety* levels were raised above those in the *None* condition. This means that intervening in the *Moral Courage* Situation further than just raising attention made participants more anxious than not intervening at all. In the *Help-Giving* Situation, all *Intervention* showed lower levels of reported *anxiety* than without *Intervention* (*None*).

Table 1: Linear Mixed Models Predicting *anxiety* and *worrying* with *Situation* and *Intervention*

Variables	*Dependent variable:* *anxiety* (1)	*worrying* (2)
(Intercept)	3.96*** (0.49)	3.46*** (0.47)
Main Effects		
Situation (Ref: *Help-Giving*)	1.63** (0.52)	2.67*** (0.51)
Attention (Ref: *None*)	−1.50* (0.70)	−1.46* (0.66)
Interactions		
Situation x *Attention*	0.83 (0.75)	0.17 (0.72)
Situation x *Awareness*	1.87* (0.75)	0.46 (0.72)
Situation x *Responsibility*	1.57* (0.74)	0.93 (0.72)
Situation x *Skills*	1.81* (0.74)	1.13 (0.72)
Situation x *Final Decision*	1.17 (0.74)	0.41 (0.72)
$R^2 conditional$	0.59	0.57
$R^2 marginal$	0.28	0.24
Note:	*p<0.05; **p<0.01; ***p<0.001	

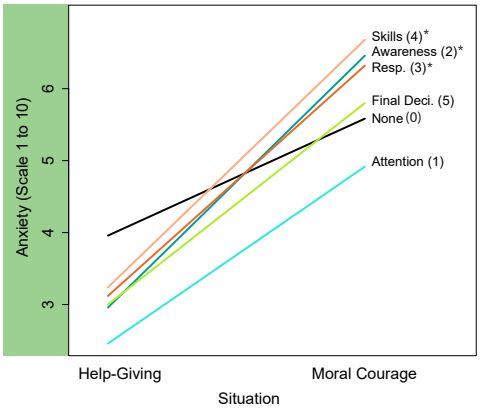
The following results are measurements taken after participants imagined the situation being resolved. Running a LMM to predict *agency* with *Situation* and *Intervention* and the participants as random effect showed no significant main or interaction effects. Further LMMs predict *infl. on help. decision* and *main reason* with *Situation* and *Intervention* (and the participant as a random effect) showed significant main effects of the levels of *Intervention* compared to *None* as seen in Table 2. This means that participants attributed their decision to help significantly more to the AR device when the device displayed *Awareness* (*M*=2.92,*SD*=1.99), *Responsibility* (*M*=2.52,*SD*=1.68), *Skills* (*M*=2.96,*SD*=1.98), and *Final Decision* (*M*=2.86,*SD*=1.86) compared to having no *Intervention* (*M*=1.54,*SD*=1.13). Regarding the *Intervention*, the findings mean that participants agreed to the statement of the AR being the main reason for helping more when the device displayed *Responsibility* (*M*=2.22,*SD*=1.52), *Skills* (*M*=2.54,*SD*=1.80), and *Final Decision* (*M*=2.50,*SD*=1.69) compared to having *Intervention* (*M*=1.46,*SD*=1.05).

Table 2: Linear Mixed Models Predicting *infl. on help. decision* and *main reason* with *Situation* and *Intervention*

	Dependent variable: *infl. on help. decision* (1)	*main reason* (2)
(Intercept)	1.46*** (0.36)	1.42*** (0.31)
Situation [Ref: *Help-Giving*]	0.17 (0.36)	0.08 (0.25)
Intervention [Ref: None]		
(1) *Attention*	0.29 (0.50)	0.08 (0.43)
(2) *Awareness*	1.29* (0.50)	0.62 (0.43)
(3) *Responsibility*	1.18* (0.50)	0.86* (0.43)
(4) *Skills*	1.50** (0.50)	1.02* (0.43)
(5) *Final Decision*	1.22* (0.50)	0.94* (0.43)
$R^2 conditional$	0.54	0.69
$R^2 marginal$	0.09	0.08
Note:	*p<0.05; **p<0.01; ***p<0.001	

5.2 *RQ2*: How does augmentation influence the perception of helping behaviour and the device's impact on it?

In *RQ2*, we were interested in how the perception of the situation and the device's augmentations were differing between *Helper* and *PiN*. Participants could not answer questions about augmentations concerning the *PiN* when there were none. We, therefore, excluded the *None* condition from the data. We also did not want to focus on specific *Intervention* but on the *Intervention*, in general, the following LMMs therefore only use *Role* and *Situation* while including participants as a random effect. In this way, we executed LMMs predicting



Anxiety		Worrying	
Help-Giving	Moral Courage	Help-Giving	Moral Courage
M (SD)	M (SD)	M (SD)	M (SD)
3.96 (2.54)	5.58 (2.59)	3.46 (2.19)	6.13 (2.25)
2.46 (2.06)	4.92 (2.72)	2.00 (1.79)	4.83 (2.73)
2.96 (2.10)	6.46 (2.21)	2.92 (2.12)	6.04 (2.51)
3.12 (2.59)	6.32 (2.30)	2.48 (2.38)	6.08 (2.27)
3.24 (2.05)	6.68 (2.39)	2.88 (1.94)	6.68 (2.27)
3.00 (2.20)	5.80 (3.04)	2.48 (1.45)	5.56 (3.03)

Figure 5: Figure (left) depicting the interaction effect between *Situation* and *Intervention*. The y-axis has been abbreviated to allow easier observation of the cross-over effect described in section 5. The table (right) shows the mean (M) and standard deviation (SD) for each combination of *Situation* and *Intervention*.

the values regarding *RQ2* (*gratitude*, *main reason*, without the device, and *SoA*) and *RQ3* (*comfort*). The models can be seen in Table 3 and will be described in the following. All were answered on a 7-point Likert scale.

For all values regarding *RQ2* we found statistically significant main effects for the two *Roles*. Participants in the *PiN* condition reported being statistically significantly (t(482) = 4.49, p<.001) more *gratitude* towards the AR device (M=3.59, SD=1.91) than participants of the *Helper* condition suspected (M=2.60, SD=1.59). Also, participants in the *PiN* condition were statistically significantly (t(482) = -6.03, p < .001) less certain that they still *would have helped* have been provided help without the Intervention (M=4.96, SD=1.68) than in the *Helper* condition (M=5.93, SD=1.51). Participants in the *PiN* agreed statistically significantly (t(482) = 4.25, p<.001) more to *main reason* for being helped being the AR device (M=3.07, SD=1.70) than in the *Helper* condition (M=2.27, SD=1.58). *PiN* also attributed a statistically significant (t(482) = 5.29, p < .001) higher level of *infl. on help. decision* (M=3.84,SD=1.91) than those in the *Helper* condition (M=2.65,SD=1.84). Finally, *PiN* attributed statistically significant (t(482) = -3.88, p<.001) lower level of *agency* (M=5.33,SD=1.48) to the *Helper* than they did themselves (M=6.00,SD=1.11).

We also found a statistically significant interaction effect between *Situation* and *Role* regarding participants thinking that the *Helper* still *would have helped* without the Intervention. It is visually depicted in Figure 6 on the left. This interaction effect means, that the *Helper* had a significant higher difference in reported values (delta = 0.76) between *Help-Giving* (M=6.31,SD=1.28) and *Moral Courage* (M=5.55,SD=1.64) condition than the *PiN* (delta = 0.26) had between *Help-Giving* (M=5.09,SD=1.71) and *Moral Courage* (M=4.83,SD=1.65) condition. Meaning that while *Helper* was more confident that they still would have helped in the *Help-Giving* condition they became less confident for the *Moral Courage* condition and got closer to the lower estimates of the *PiN*.

5.3 Comfort with the Augmentations

Regarding reported the reported levels of *comfort*, we found that participants felt generally statistically significant (t(482) = -6.37, p<0.001) less comfortable with *Interventions* in the *Moral Courage* (M=3.48, SD=1.84) than in the *Help-Giving* (M=4.16, SD=1.97) condition. Also, *PiN* felt significantly (t(482) = -2.97, p<0.01) less comfortable (M=3.63, SD=1.98) with the *Interventions* than a *Helper* (M=4.01, SD=1.87).

The model also revealed an interaction effect for *comfort* between *Role* and *Situation* (t(482) = 2.96, p<0.003). While in the *Help-Giving* Situation *Helper* (M=4.52,SD=1.87) and *PiN* (M=3.80,SD=2.00) had a higher gap between their *comfort* ratings (delta = 0.72), in the *Moral Courage* Situation the difference between *Helper* (M=3.50,SD=1.72) and *PiN* (M=3.45,SD=1.96) gets less substantial (delta = 0.05). This interaction can be seen in Figure 5,c.

Figure 6 displays the mean ratings of *comfort* split by *Situation* and *Role* as well as *Intervention*. Here mean values higher than the neutral of 3.5 (on a 7-point Likert-Scale) are highlighted in a darker green. It can be observed that there are only five combinations of *Intervention* and *Situation* in which both *Helper* and *PiN* felt comfortable with the augmentation. This was *Awareness* in the *Help-Giving* condition as well as *Responsibility* and *Skills* in both *Situations*.

6 DISCUSSION

The following section will discuss the results of our study. Here, we draw conclusions on how augmenting social situations can influence potential help-giving behaviour and the social perception of this situation. We also conclude how prosocial augmentation and augmentation, in general, could influence social relations and discuss ethical implications.

Table 3: Linear Mixed Models Predicting *anxiety* and *worrying* with *Situation* and *Intervention*

	Dependent variables:					
	gratitude (towards AR device)	*would have helped* (without AR device)	*main reason* (for helping was AR device)	*infl. on help. decision* (of AR device)	agency (of *Helper*)	*comfort* (with augmentations)
(Intercept)	2.47*** (0.16)	6.31*** (0.14)	2.13*** (0.15)	2.56*** (0.17)	6.03*** (0.12)	4.52*** (0.17)
Main Effects						
Situation [Ref: H.-Giving]	0.27* (0.12)	−0.76*** (0.15)	0.28* (0.12)	0.18 (0.16)	−0.06 (0.08)	−1.02*** (0.16)
Role [Ref: *Helper*]	1.01*** (0.22)	−1.22*** (0.20)	0.89*** (0.21)	1.27*** (0.24)	−0.65*** (0.17)	−0.72** (0.24)
Interaction						
Situation x *Role*	−0.05 (0.18)	0.49* (0.21)	−0.19 (0.17)	−0.16 (0.22)	−0.03 (0.12)	0.68** (0.23)
R^2*conditional*	0.71	0.54	0.69	0.61	0.77	0.58
R^2*marginal*	0.08	0.11	0.06	0.09	0.06	0.05
Note:	*p<0.05; **p<0.01; ***p<0.001					

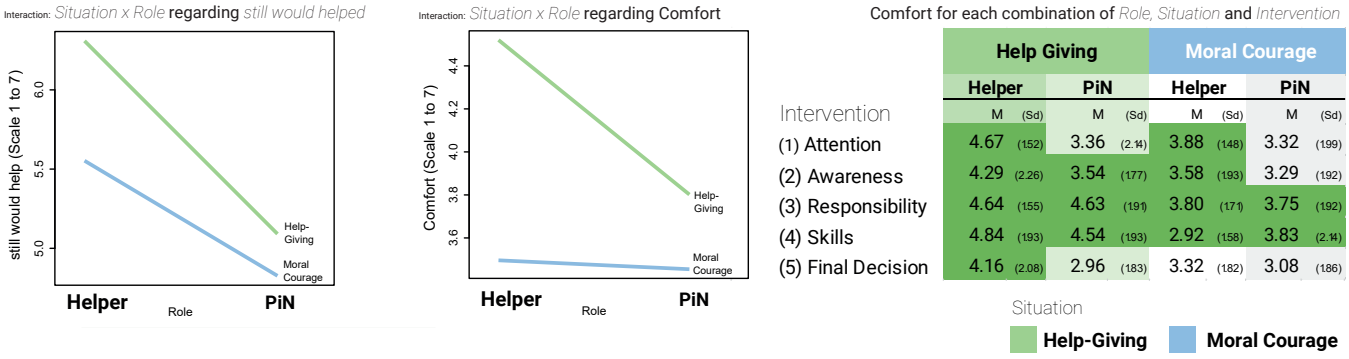


Figure 6: A figure (left) depicting the interaction between *Role* and *Situation* regarding *would have helped*. Figure (middle) depicts the interaction between *Role* and *Situation* regarding *comfort*. Please be aware that the axis scaling between the figures differs. Table (right) including the ratings of *comfort* split by *Situation* and *Role* as well as *Intervention*. Means equal to or above the neutral point of 3.5 are highlighted in a darker green, indicating that participants did not report feeling uncomfortable.

6.1 Can AR systems in potential helping situations influence the (1) Situation Assessment and (2) Feelings induced by the situation?

Contrary to what was suspected in *RQ1* (see section 3), the five augmentation levels did not significantly impact a potential helper's Situation Assessment. This means that in neither augmentation, we found a significant influence on *Pluralistic Ignorance*, *Evaluation Apprehension*, and *Diffusion of Responsibility*. This might originate because even though we tried introducing the bystander effect in the video (see Figure 4.2), the camera's perspective firmly focused on the particular situation. This emphasis on the particular helping opportunity might have overemphasized what was expected from the participant (helping). This focus is also reflected in the overall willingness to help even without any augmentation (*Help-Giving*: *M*=5.66 and *Moral Courage* *M*=4.29; on a 7-point scale). As related work has found that being with a known person and seeing their reactions impacts the bystander effect [52], the novelty of just being introduced to the device could have influenced the effect of the augmentations.

Influence on Anxiety. Regarding the feelings towards the situation, we found a significant influence on the reported anxiety levels. Here, we found that just making the user aware of the situation (*Attention*) significantly influenced the potential helpers to feel less anxiety. We also found a significant cross-over interaction when looking at augmentations of different situations. Participants generally reported lower anxiety levels when the *Help-Giving* situation was augmented compared to no augmentation. However, this was reversed for augmentations containing textual descriptions in the *Moral Courage* situation, as describing the situation led to a significant rise in anxiety.

The difference between situations might be explained by Osswald et al. [65]'s distinguishing between moral courage and helping-giving regarding expected social consequences. In a *Help-Giving*

situation, a potential helper can expect mostly positive social consequences. They had an entity that also interpreted the situation, telling them about the righteousness of their helping endeavour., leading to a potential increase in expectation of possible positive social consequences (e.g., being thanked by the *PiN*). Additionally, the user might use the AR device as a scapegoat if the attempt to help is met with discontent. However, in the *Moral Courage* condition, the AR glasses cannot prevent potential consequences that might arise from interfering (physical confrontation).

What this means: Highlighting helping behaviour situations without textually defining them can make users feel less anxious about the situation happening around them. In turn, the context defines the influences of providing text, which can even lead to higher anxiety levels. As pointed out by related work, anxiety and fear about the situation can hinder helping behaviour [29, 47, 65].

Attribution of Decision. Regarding the attribution of the helping behaviour, we found evidence of the augmentation's influence. While we found that *Helper* generally disagreed with the fact that the AR device influenced their decision process, all textual interventions (*Awareness* to *Final Decision*) led to significantly higher levels of reported perceived influence of the device. We found similar results regarding the AR device's attribution as the main reason for the helping behaviour. Here *Responsibility*, *Skills*, and *Final Decision* lead to a significantly higher attribution towards the AR device being the main reason for the helper to intervene.

This significant increase in perceived influence and "being the main reason to help" indicates that the AR device was perceived as an entity that had impacted the situation. While the helper still attributed the majority of their decision to help to themselves, introducing the augmentation significantly lowered their confidence in being the main reason.

6.2 *How does augmentation influence the perception of helping behaviour and the device's impact on it?*

Previous work has shown that the comfort felt by the person perceiving an augmentation and the person being augmented can significantly differ [71, 72]. This creates an inherent asymmetry in the perception of the augmentation and the perception of the AR device between the wearer and augmentation target (here: *Helper* and *PiN*). While our findings replicated this asymmetry of comfort in *Helper* and *PiN*, we also found this asymmetry in other measured variables. Our findings around the authority of the helping decision (e.g., "what was the main reason to help") and responsibility (e.g., "who to be thankful to") show that the asymmetry of perception in AR can impact more aspects of our social structure.

We found that *Helper* generally disagreed with the fact that the AR device influenced their decision process, significantly different from the perception of the *PiN*. They also showed significantly more gratitude towards the AR device than the *Helper* suspected and could significantly less strongly imagine that the *Helper* would have helped even without AR device. Finally, the *PiN* agreed significantly more to the Augmented Reality device being the main reason they were helped than the helper.

These findings indicate that the *PiN* considered the impact of the device stronger than the helper. Like previous work has shown that the augmentations can evoke different levels of comfort depending on the actor's role, we also find this perception gap in the perceived influences on helping behaviour. This demonstrates that the asymmetry of perception of interactions, including AR, is more complex than only around feeling-centered metrics such as comfort. It also includes the perception of authority and responsibility for a person's behaviour. Here, we want to emphasize that this asymmetry did not arise from an actual information asymmetry (both helper and helped were seeing the same demonstration videos)

In his seminal work "Augmenting the Human Intellect", Douglas Engelbart [18] presents a conceptual framework and a vision of how technology can enhance humans' abilities to solve complex problems. He provides examples: *more-rapid comprehension, better comprehension, the possibility of gaining a useful degree of comprehension in a previously too complex situation, and the possibility of finding solutions to problems that before seemed insoluble.* Augmented Reality in helping situations is a fitting example demonstrating how technology could help us with "better comprehension in a complex situation". However, our findings around the disagreement of authority (why was helped) and responsibility (whom should I be grateful to) between helper and person in need point to a significant problem that might arise from such human augmentation. Instead of perceiving the user as one with the technology, *Helper* and *PiN* perceived the AR device as a third entity that impacts the helper's behaviour. Additionally, the perception was asymmetric when asked how it impacted the decision. This disagreement has potentially severe consequences on our social structure, built around the individual's autonomy and independence (free will even). For example, when one person performs a good act, they take full responsibility and credit for the action. However, when the same act is performed while augmented, the helper and person in need disagree on how much the technology impacted the decision. Whom should we be thankful to? Did the counterpart only perform this action because the technology asked them to? These questions can potentially impact our interpersonal interaction once AR becomes widely distributed.

In an ideal scenario of human augmentation, the technology and the user are forming a symbiotic entity [55]. However, our findings unveil an essential problem in social interactions that might arise from this symbiosis around the perception of others in complex social situations. We found that the technology could become an individual entity that is interfering with how others perceive the augmented person. Instead of focusing exclusively on optimizing interaction and usability of the augmented person, we argue that future research should start exploring how we could design augmentations that other people in social interactions perceive as a symbiotic part of the user and not as a third entity, steering and interfering with decisions and actions.

6.3 Ethical Considerations

To build AR technology where AR devices can be socially acceptable, future AR devices and their applications should be designed with every involved person's comfort in mind [71]. In our study, we only found five combinations of *Situation* and augmentations in which all involved did not feel uncomfortable with the situation (see Figure 6).

This means that only augmenting *Help-Giving* situations with *Awareness*, *Responsibility*, and *Final Decision* makes everyone involved feel comfortable. In a *Moral Courage* situation, only *Responsibility* and *Skills* should be augmented to ensure everyone's comfort.

While comfort is essential, considerations may occur when a positive trade-off matches discomfort. As discussed, highlighting the situation and, therefore, the *PiN* helped the potential helper feel less anxious about the situation. However, this resulted in the *PiN*, in turn, feeling less comfortable. This demonstrates well that optimizing AR applications in the future will always need to consider the perspectives of everyone involved, which might even result in a contradiction/trade-off. Once AR becomes widely distributed, application developers must create new metrics and decide individual trade-offs about specific augmentations. For example, would the reduction of anxiety in the helper warrant the creation of discomfort for the person in need? While we cannot answer this question, our findings emphasize the severity of the asymmetry in perception that future AR augmentations will create. Comfort and anxiety are just one example of an asymmetry trad-off in AR, which emphasizes the importance of considering the needs of all participants involved in an augmentation.

On another note, similar to discussions about the liability of accidents in autonomous cars [32, 58, 76], such technology could also lead to users getting in danger or even getting hurt. In our study, the proposed actions were based on an official guide by the German police. On the one hand, giving the right instructions could lead to a person de-escalating the situation without any further harm done. On the other hand, even executing the theoretical right steps in a moral courage situation to de-escalate could fail and lead to injury. Who will be liable in the second case? Moreover, analogous to the question Marchant and Lindor [58] asked about accidents in autonomous cars, what weight will the courts give to the overall comparative safety that such systems could provide when determining whether those involved in harm should be held liable?

Another question is raised about the interpretation of the situation itself. Even when used in good faith, the automated interpretation of a situation is predestined to carry biases. The system interpreting the situation bases it on how it was programmed or what it learned from the set it was trained on. These sets, in turn, can contain implicit biases and therefore reflect the moral and interpretation of its creators or that embedded in the data it is based on [16, 64]. If regulated by authorities, this might, e.g., reflect the moral principles of a democratic society or those of an authoritarian leadership. If not regulated, it will reflect those of its developers or recreate societal biases embedded in data. With AR becoming a part of daily life, users wearing a device able to interpret situations are, therefore, exposed to and influenced by the interpretations and the moral views embedded. As those algorithms "find patterns within data-sets that reflect implicit biases and, in so doing, emphasize and reinforce these biases as global truth" [35], these reinforcements could be carried on to the users.

We argue that while computer scientists and tech companies might provide the technology to enable such features, a decision to do so should not be rushed. Instead, a public discussion about trade-offs, liability, and the reinforcement of biases has to be held to mitigate potential weighty consequences.

7 LIMITATIONS

Researchers have identified gaps between what they intend to do and what they actually do [10]. This is the so-called Intention–Behaviour Gap [33]. Not having participants act but state their intentions in an imagined situation might, therefore, not reflect their actual behaviour in the same situation. Analogous to Levine et al. [53], we do not try to measure actual behaviour but the influences on the perception of the situation and feelings as well as the mere intent to act.

Participants in our study were not put in a situation where they faced danger. Instead, they were presented with a video representation and asked to imagine being part of it. While one might expect that actual exposure to the danger of such a situation (or the mere adding of additional stimuli like sound) can be expected to yield stronger responses, we, nevertheless, can observe that the moral courage condition led to significantly higher levels of reported anxiety and worrying. We can, therefore, reason that the participants visually and imaginatively experienced the danger of the moral courage situation.

Participants were overall exposed to a 1 min video vignette (50-sec introduction + 11-sec condition) representing the situation. While we argue that this exposure was long enough to enable participants to imagine themselves in it, a longer exposure might have yielded higher levels of immersion.

Also, to allow participants in the *PiN* condition to understand the augmentations, we also had to show them the situation from the perspective of the *Helper*. While necessary to explore the augmentations, it might have made it more difficult for participants in the *PiN*'s condition to imagine themselves being in the situation.

The high difference between $R^2 conditional$ and $R^2 marginal$, as well as the high standard deviations, show that it is highly individual and a high portion of the effect can be accounted towards individual differences. While the fixed effects had an influence, individual differences also played a significant part.

While imagining being in a particular situation is an established method in social acceptability research [41], in our work, we did not measure reactions to the technology, augmentations, and helping situations but representations of them. This, in addition to novelty factors of AR technology, might impact our findings.

As previous work has shown that a person's physical attributes like perceived sex [48] can influence the helping reaction towards them, our results might not be generalizable to persons with different physical appearances. While one bystander alone can produce a bystander effect Latané and Rodin [52], adding more bystanders could have had a more considerable impact Latané and Dabbs Jr [49].

8 FUTURE DIRECTIONS

In our work, we explored how AR could influence the perception of prosocial *Help-Giving* and *Moral Courage* situations. Future work should explore other prosocial decisions like the decision to participate in "sharing, comforting, donating, or volunteering" [p.1 17] This would allow a more comprehensive picture of the impact of AR on us as social beings and our social structure. Also, with the ongoing development of new wearable AR technology, our study should eventually become feasible to recreate in a real-life context. Recreating the study would enable tackling the question of whether augmentation can not only influence how the situation is perceived but also influence actual action.

9 CONCLUSION

With wearable AR technology, a new dynamic layer of information can be added to our sensory repertoire able to help by interpreting their surroundings with them [15, 56]. Future AR devices will not only be used in isolation but might also find application in the social situations that occur in a user's daily life [31, 68, 71, 72]. While prior work has already established the concept of augmenting an interlocutor with personal information [1, 1, 28, 44], future wearable AR devices could also make users aware of social situations happening in their surroundings and guiding them through those. This work raises the question of how such interventions could look, how they can influence the perception of a situation, and how they would impact how social situations and the augmented entity are perceived. To explore this, we look at one highly discussed prosocial behaviour: helping behaviour in daily life. In this context, we create mock-up videos of a help-giving situation (a person not getting through a door) and a situation needing moral courage (a person being attacked). In a mixed factorial online video experiment (n=294), we explore how intervening in a potential helper's decision process on five levels influences how both helper and person in need feel about the situation and potential help. We explore this in both the moral courage and help-giving situation while getting first hints on how such an intervention could influence the assessment of and feelings towards a situation. While we found that augmentations did not influence situational assessment in regard of bystander-effect-related psychological processes and reported intent to help, we found influences on situation-related anxiety and differences in how much helpers attributed their decision toward their AR device. We also found that, like in previous work, the augmented person feels less comfortable than the device's wearer. We discuss how these differences could collide with other desired outcomes of an AR application. This work also reveals that an AR device can be perceived as an influential actor in a helping situation. How much influence is attested herein varies with the role, with the helped person attesting a greater influence. We further discuss how this diffusion in attribution could impact social relationships. Therefore, our work first sheds light on how the interpretation and successive augmentation of helping behaviour through AR could impact the perception of helping behaviour.

ACKNOWLEDGMENTS

We want to thank Michael Schelling, who, without hesitation, agreed to lend his exceptional acting skills to the making of this paper. We also want to thank Sven-Ole and Claudia for always being there. This work was supported by the Landesgraduiertenförderung (LGFG) Scholarship for PhD students.

REFERENCES

[1] Alessandro Acquisti, Ralph Gross, and Frederic D Stutzman. 2014. Face recognition and privacy in the age of augmented reality. *Journal of Privacy and Confidentiality* 6, 2 (2014), 1.

[2] Francis M Adams and Charles E Osgood. 1973. A cross-cultural study of the affective meanings of color. *Journal of cross-cultural psychology* 4, 2 (1973), 135–156.

[3] Herman Aguinis and Kyle J Bradley. 2014. Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational research methods* 17, 4 (2014), 351–371.

[4] Ahmed Alnabhan and Brian Tomaszewski. 2014. INSAR: Indoor navigation system using augmented reality. In *Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*. ACM, 36–43.

[5] Christiane Atzmüller and Peter M Steiner. 2010. Experimental vignette studies in survey research. *Methodology* (2010).

[6] Amanda Baker, Naomi Simon, Aparna Keshaviah, Amy Farabaugh, Thilo Deckersbach, John J Worthington, Elizabeth Hoge, Maurizio Fava, and Mark P Pollack. 2019. Anxiety Symptoms Questionnaire (ASQ): development and validation. *General psychiatry* 32, 6 (2019).

[7] Carlos Bermejo, Zhanpeng Huang, Tristan Braud, and Pan Hui. 2017. When augmented reality meets big data. In *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*. IEEE, IEEE, New York, NY, USA, 169–174.

[8] Leonard Bickman. 1972. Social Influence and Diffusion of Responsibility in an Emergency. *Journal of Experimental Social Psychology* 8, 5 (1972), 438–445.

[9] Arthur P Brief and Stephan J Motowidlo. 1986. Prosocial organizational behaviors. *Academy of management Review* 11, 4 (1986), 710–725.

[10] Michal J Carrington, Benjamin A Neville, and Gregory J Whitwell. 2010. Why ethical consumers don't walk their talk: Towards a framework for understanding the gap between the ethical purchase intentions and actual buying behaviour of ethically minded consumers. *Journal of business ethics* 97, 1 (2010), 139–158.

[11] Chia-Yen Chen, Bao Rong Chang, and Po-Sen Huang. 2014. Multimedia augmented reality information system for museum guidance. *Personal and ubiquitous computing* 18, 2 (2014), 315–322.

[12] Yi Fei Cheng, Hang Yin, Yukang Yan, Jan Gugenheimer, and David Lindlbauer. 2022. Towards Understanding Diminished Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 549, 16 pages. https://doi.org/10.1145/3491102.3517452

[13] Russell D. Clark and Larry E. Word. 1972. Why Don't Bystanders Help? Because of Ambiguity? *Journal of Personality and Social Psychology* 24, 3 (1972), 392.

[14] Ian Cunnings. 2012. An overview of mixed-effects statistical models for second language researchers. *Second Language Research* 28, 3 (2012), 369–382.

[15] Jena Daniels, Jessey N Schwartz, Catalin Voss, Nick Haber, Azar Fazel, Aaron Kline, Peter Washington, Carl Feinstein, Terry Winograd, and Dennis P Wall. 2018. Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. *NPJ digital medicine* 1, 1 (2018), 1–10.

[16] Ketki V. Deshpande, Shimei Pan, and James R. Foulds. 2020. Mitigating Demographic Bias in AI-Based Resume Filtering. In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization* (Genoa, Italy) *(UMAP '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 268–275. https://doi.org/10.1145/3386392.3399569

[17] John F. Dovidio and Jillian C. Banfield. 2015. Prosocial Behavior and Empathy. In *International Encyclopedia of the Social and Behavioral Sciences (Second Edition)* (second edition ed.), James D. Wright (Ed.). Elsevier, Oxford, 216–220. https://doi.org/10.1016/B978-0-08-097086-8.24024-5

[18] Douglas C Engelbart. 1962. Augmenting human intellect: A conceptual framework. *Menlo Park, CA* (1962), 21.

[19] Alberto N Escalante-B and Laurenz Wiskott. 2020. Improved graph-based SFA: Information preservation complements the slowness principle. *Machine Learning* 109, 5 (2020), 999–1037.

[20] KONSTANTINOS P Ferentinos and MYRTO S Barda. 2021. A deep learning plants identification model for augmented reality touring in urban parks. In *Proceedings of the CIGR 5th International Conference, Quebec City, QC, Canada*. 11–14.

[21] Peter Fischer, Tobias Greitemeyer, Fabian Pollozek, and Dieter Frey. 2006. The Unresponsive Bystander: Are Bystanders More Responsive in Dangerous Emergencies? *European journal of social psychology* 36, 2 (2006), 267–278.

[22] P Fischer, T Greitemeyer, S Schulz-Hardt, D Frey, E Jonas, and T Rudukha. 2004. Zivilcourage and helping behavior: The impact of negative social consequences on the perception of prosocial behavior. *Zeitschrift fur Sozialpsychologie* 35, 2 (2004), 61–66.

[23] Peter Fischer, Joachim I. Krueger, Tobias Greitemeyer, Claudia Vogrincic, Andreas Kastenmüller, Dieter Frey, Moritz Heene, Magdalena Wicher, and Martina Kainbacher. 2011. The Bystander-Effect: A Meta-Analytic Review on Bystander Intervention in Dangerous and Non-Dangerous Emergencies. *Psychological bulletin* 137, 4 (2011), 517.

[24] Peter Fischer, Joachim I Krueger, Tobias Greitemeyer, Claudia Vogrincic, Andreas Kastenmüller, Dieter Frey, Moritz Heene, Magdalena Wicher, and Martina Kainbacher. 2011. The bystander-effect: a meta-analytic review on bystander intervention in dangerous and non-dangerous emergencies. *Psychological bulletin* 137, 4 (2011), 517.

[25] Mark G Frank and Thomas Gilovich. 1988. The dark side of self-and social perception: black uniforms and aggression in professional sports. *Journal of personality and social psychology* 54, 1 (1988), 74.

[26] Siyao Fu, Haibo He, and Zeng-Guang Hou. 2014. Learning race from face: A survey. *IEEE transactions on pattern analysis and machine intelligence* 36, 12 (2014), 2483–2509.

[27] Edward Gibson, Steve Piantadosi, and Kristina Fedorenko. 2011. Using Mechanical Turk to obtain and analyze English acceptability judgments. *Language and Linguistics Compass* 5, 8 (2011), 509–524.

[28] Amin Golnari, Hossein Khosravi, and Saeid Sanei. 2020. DeepFaceAR: Deep Face Recognition and Displaying Personal Information via Augmented Reality. In *2020 International Conference on Machine Vision and Image Processing (MVIP)*. IEEE, IEEE, New York, NY, USA, 1–7.

[29] Tobias Greitemeyer, Peter Fischer, Andreas Kastenmüller, and Dieter Frey. 2006. Civil Courage and Helping Behavior: Differences and Similarities. *European Psychologist* 11, 2 (2006), 90–98.

[30] Tobias Greitemeyer, Silvia Osswald, Peter Fischer, and Dieter Frey. 2007. Civil Courage: Implicit Theories, Related Concepts, and Measurement. *The Journal of Positive Psychology* 2, 2 (2007), 115–119.

[31] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[32] Jeffrey K Gurney. 2013. Sue my car not me: Products liability and accidents involving autonomous vehicles. *U. Ill. JL Tech. & Pol'y* (2013), 247.

[33] Louise M Hassan, Edward Shiu, and Deirdre Shaw. 2016. Who says there is an intention–behaviour gap? Assessing the empirical evidence of an intention–behaviour gap in ethical consumption. *Journal of Business Ethics* 136, 2 (2016), 219–236.

[34] Ramón Hervás, José Bravo, and Jesús Fontecha. 2013. An assistive navigation system based on augmented reality and context awareness for people with mild cognitive impairments. *IEEE Journal of Biomedical and Health Informatics* 18, 1 (2013), 368–374.

[35] Ayanna Howard and Jason Borenstein. 2018. The ugly truth about ourselves and our robot creations: the problem of bias and social inequity. *Science and engineering ethics* 24, 5 (2018), 1521–1536.

[36] Rhidian Hughes and Meg Huby. 2002. The application of vignettes in social and nursing research. *Journal of advanced nursing* 37, 4 (2002), 382–386.

[37] Ted L Huston, Mary Ruggiero, Ross Conner, and Gilbert Geis. 1981. Bystander intervention into crime: A study based on naturally-occurring episodes. *Social Psychology Quarterly* (1981), 14–23.

[38] Daniela Niesta Kayser, Tobias Greitemeyer, Peter Fischer, and Dieter Frey. 2010. Why Mood Affects Help Giving, but Not Moral Courage: Comparing Two Types of Prosocial Behaviour. *European Journal of Social Psychology* 40, 7 (2010), 1136–1157. https://doi.org/10.1002/ejsp.717

[39] Susan Herrington Kidd and Helen Crompton. 2016. Augmented learning with augmented reality. In *Mobile learning design*. Springer, 97–108.

[40] Suna Pirita Kinnunen and Sabine Windmann. 2013. Dual-Processing Altruism. *Frontiers in Psychology* 4 (2013). https://doi.org/10.3389/fpsyg.2013.00193

[41] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. 2020. Social acceptability in HCI: A survey of methods, measures, and design strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–19.

[42] Marion Koelle, Abdallah El Ali, Vanessa Cobus, Wilko Heuten, and Susanne CJ Boll. 2017. All about Acceptability? Identifying Factors for the Adoption of Data Glasses. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 295–300. https://doi.org/10.1145/3025453.3025749

[43] Marion Koelle, Matthias Kranz, and Andreas Möller. 2015. Don't Look at Me That Way! Understanding User Attitudes Towards Data Glasses Usage. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 362–372. https://doi.org/10.1145/2785830.2785842

[44] Martin Kurze and Axel Roselius. 2011. Smart Glasses Linking Real Live and Social Network's Contacts by Face Recognition. In *Proceedings of the 2nd Augmented Human International Conference* (Tokyo, Japan) *(AH '11)*. Association for Computing Machinery, New York, NY, USA, Article 31, 2 pages. https://doi.org/10.1145/1959826.1959857

[45] Mikko Kytö and David McGookin. 2017. Augmenting Multi-Party Face-to-Face Interactions Amongst Strangers with User Generated Content. *Computer Supported Cooperative Work (CSCW)* 26, 4 (2017), 527–562.

[46] Mikko Kytö and David McGookin. 2017. Investigating user generated presentations of self in face-to-face interaction between strangers. *International Journal of Human-Computer Studies* 104 (2017), 1–15.

[47] Vicki D. Lachman. 2007. Moral Courage: A Virtue in Need of Development? *Medsurg nursing* 16, 2 (2007), 131–133.

[48] Mary R Laner, Mary H Benin, and Nicole A Ventrone. 2001. Bystander attitudes toward victims of violence: Who's worth helping? *Deviant Behavior* 22, 1 (2001), 23–42.

[49] Bibb Latané and James M Dabbs Jr. 1975. Sex, group size and helping in three cities. *Sociometry* (1975), 180–194.

[50] Bibb Latané and John M. Darley. 1970. *The Unresponsive Bystander: Why Doesn't He Help?* Appleton-Century-Crofts.

[51] Bibb Latané and Steve Nida. 1981. Ten Years of Research on Group Size and Helping. *Psychological bulletin* 89, 2 (1981), 308.

[52] Bibb Latané and Judith Rodin. 1969. A lady in distress: Inhibiting effects of friends and strangers on bystander intervention. *Journal of Experimental Social Psychology* 5, 2 (1969), 189–202.

[53] Mark Levine, Clare Cassidy, Gemma Brazier, and Stephen Reicher. 2002. Self-categorization and bystander non-intervention: Two experimental studies 1. *Journal of Applied Social Psychology* 32, 7 (2002), 1452–1463.

[54] Xiaobai Li, Xiaopeng Hong, Antti Moilanen, Xiaohua Huang, Tomas Pfister, Guoying Zhao, and Matti Pietikäinen. 2017. Towards reading hidden emotions: A comparative study of spontaneous micro-expression spotting and recognition methods. *IEEE transactions on affective computing* 9, 4 (2017), 563–577.

[55] Joseph CR Licklider. 1960. Man-computer symbiosis. *IRE transactions on human factors in electronics* 1 (1960), 4–11.

[56] Runpeng Liu, Joseph P Salisbury, Arshya Vahabzadeh, and Ned T Sahin. 2017. Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching. *Frontiers in pediatrics* 5 (2017), 145.

[57] Xiao Ma, Jeff Hancock, and Mor Naaman. 2016. Anonymity, intimacy and self-disclosure in social media. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. ACM, New York, NY, USA, 3857–3869.

[58] Gary E Marchant and Rachel A Lindor. 2012. The coming collision between autonomous vehicles and the liability system. *Santa Clara L. Rev.* 52 (2012), 1321.

[59] Joseph F McCarthy, David W McDonald, Suzanne Soroczak, David H Nguyen, and Al M Rashid. 2004. Augmenting the social space of an academic conference. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*. ACM, New York, NY, USA, 39–48.

[60] David McGookin and Mikko Kytö. 2016. Understanding User Attitudes to Augmenting Face-to-Face Interactions with Digital and Social Media. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (Rovaniemi, Finland) *(MUM '16)*. Association for Computing Machinery, New York, NY, USA, 285–296. https://doi.org/10.1145/3012709.3012731

[61] Alessandro Mulloni, Hartmut Seichter, and Dieter Schmalstieg. 2011. Handheld Augmented Reality Indoor Navigation with Activity-Based Instructions. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden) *(MobileHCI '11)*. Association for Computing Machinery, New York, NY, USA, 211–220. https://doi.org/10.1145/2037373.2037406

[62] Tien T. Nguyen, Duyen T. Nguyen, Shamsi T. Iqbal, and Eyal Ofek. 2015. The Known Stranger: Supporting Conversations between Strangers with Personalized Topic Suggestions. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 555–564. https://doi.org/10.1145/2702123.2702411

[63] Geoff Norman. 2010. Likert scales, levels of measurement and the "laws" of statistics. *Advances in health sciences education* 15, 5 (2010), 625–632.

[64] Eirini Ntoutsi, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, Franco Turini, Symeon Papadopoulos, Emmanouil Krasanakis, et al. 2020. Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 10, 3 (2020), e1356.

[65] Silvia Osswald, Dieter Frey, and Bernhard Streicher. 2011. Moral Courage. In *Justice and Conflicts*. Springer, 391–405.

[66] Jane E Palmer, Sarah C Nicksa, and Sarah McMahon. 2018. Does who you know affect how you act? The impact of relationships on bystander intervention in interpersonal violence situations. *Journal of interpersonal violence* 33, 17 (2018), 2623–2642.

[67] Levy Peter, Lundgren Diane, Ansel Marc, Fell David, Fink Betty, and JE McGrath. 1972. Bystander effect in a demand-without-threat situation. *Journal of Personality and Social Psychology* 24, 2 (1972), 166.

[68] Xukan Ran, Carter Slocum, Maria Gorlatova, and Jiasi Chen. 2019. ShareAR: Communication-Efficient Multi-User Mobile Augmented Reality. In *Proceedings of the 18th ACM Workshop on Hot Topics in Networks* (Princeton, NJ, USA) *(HotNets '19)*. Association for Computing Machinery, New York, NY, USA, 109–116. https://doi.org/10.1145/3365609.3365867

[69] Rasmus K Rendsvig. 2014. Pluralistic ignorance in the bystander effect: informational dynamics of unresponsive witnesses in situations calling for intervention. *Synthese* 191, 11 (2014), 2471–2498.

[70] Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 887–896. https://doi.org/10.1145/1753326.1753458

[71] Jan Ole Rixen, Mark Colley, Ali Askari, Jan Gugenheimer, and Enrico Rukzio. 2022. Consent in the Age of AR: Investigating The Comfort With Displaying Personal Information in Augmented Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 295, 14 pages. https://doi.org/10.1145/3491102.3502140

[72] Jan Ole Rixen, Teresa Hirzle, Mark Colley, Yannick Etzel, Enrico Rukzio, and Jan Gugenheimer. 2021. Exploring Augmented Visual Alterations in Interpersonal Communication. In *Proceedings of the 2021 CHI Conference on Human Factors in*

Computing Systems. ACM, New York, NY, USA, 1–11.

[73] Gregory K. Rutkowski, Charles L. Gruder, and Daniel Romer. 1983. Group Cohesiveness, Social Norms, and Bystander Intervention. *Journal of Personality and Social Psychology* 44, 3 (1983), 545.

[74] Stephen J Sauer. 2011. Taking the reins: the effects of new leader status and leadership style on team performance. *Journal of Applied Psychology* 96, 3 (2011), 574.

[75] Linda Zener Solomon, Henry Solomon, and Ronald Stone. 1978. Helping as a function of number of bystanders and ambiguity of emergency. *Personality and Social Psychology Bulletin* 4, 2 (1978), 318–321.

[76] Araz Taeihagh and Hazel Si Min Lim. 2019. Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport reviews* 39, 1 (2019), 103–128.

[77] Adam Tapal, Ela Oren, Reuven Dar, and Baruch Eitam. 2017. The sense of agency scale: A measure of consciously perceived control over one's mind, body, and the immediate environment. *Frontiers in psychology* 8 (2017), 1552.

[78] Yilun Wang and Michal Kosinski. 2018. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of personality and social psychology* 114, 2 (2018), 246.

[79] Qianli Xu, Shue Ching Chia, Bappaditya Mandal, Liyuan Li, Joo-Hwee Lim, Michal Akira Mukawa, and Cheston Tan. 2016. SocioGlass: social interaction assistance with face recognition on google glass. *Scientific Phone Apps and Mobile Devices* 2, 1 (2016), 1–4.