

ABSTRACT

[type=editor, auid=000,bioid=1, prefix=Sir, role=Researcher, orcid=0000-0001-7511-2910, facebook=, twitter=, linkedin=, gplus=]

ORCID(s):

The telepresence contact hypothesis

Avner Peled, Avner Peled^{a,b}

ARTICLE INFO

Keywords:
Soft robotics
Second keyword
Third keyword
Fourth keyword

The pervasive role of technology in intensifying the ability of mankind to harm one another is well known; the use of technology to promote peace at both the collective and personal levels is considerably more modest. Over the years there have been calls within the Human-Computer Interaction research community to promote the use of technology to support peace in world conflicts [64, 39]. Often when people think of a technological contribution to conflict resolution, emphasis is placed on negotiation and decision support for policy makers and nation leaders. A different question that is addressed in the current paper, is how to help reconcile the ‘common’ people in a situation of conflict and build more positive intergroup relations from the bottom up.

One of the most prominent models that acts as a guideline for this approach is the contact hypothesis [3], that states that under the right conditions, encounters with members of the opposing group (i.e., the outgroup) can lead to reduced prejudice and more harmonious intergroup relations. We propose using telepresence robots as a communication medium for such contact, combining elements from the virtual world of online interactions and from real face-to-face (FtF) encounters.

We begin by reviewing research on intergroup contact and the need a new technological medium, introducing robots as an alternative form of contact that has to our knowledge not been considered previously. We then present a conceptual framework based on previous research, outlining the conditions, considerations and hypotheses that we believe are essential for robot-mediated contact. We conclude with an agenda for research in this newly emerging field of robotic intergroup contact.

1. Intergroup contact hypothesis

The contact hypothesis as formulated by Gordon Allport in his seminal book *The Nature of Prejudice* [3] specifies four conditions that need to be fulfilled during contact: equal status, having common goals, active co-operation, and institutional support. Fifty years later, a meta-analysis validated the hypothesis across a large number of studies [120]. It was shown, however, that Allport's conditions are not strictly essential for a positive outcome, rather they are factors among others that

ABSTRACT

facilitate it. Later research focused on expanding the theory to more conditions such as forming cross-group friendships, identifying the affective drivers, such as empathy and anxiety, that play a mediating role in contact, and discussing how levels of *group salience* moderate the generalization of attitudes from the interpersonal level to the group level [121, 18]

1.2 Online contact

Communication technologies expand the models of contact and add new modalities of interaction while compromising on the benefits of traditional FtF encounters. Research on online intergroup contact has shown its potential to reduce prejudice and aid in conflict resolution [5, 52, 147]. Virtual reality has also been studied as a medium, both as a space for dialog, and as an “empathy machine” that allows intergroup members to immerse themselves in the perspective of the other side [53, 54, 73]. Nevertheless, online contact is not always constructive. The distant nature of the remote medium reduces the sense of *presence* in the social situation [46, 90], which in turn participants less accountable for their actions and less engaged in the conversation [149, 127]. The lack of non-verbal cues [23] obstructs the path to a mutual understanding and impairs the turn-taking process, which may evoke negative feelings between the group members such as anger and frustration [70].

Critical theorists have repeatedly warned against an abstraction of human nature. Terms such as *reification*, *rationalization*, and *fetishism* are used to describe the underlying condition of prejudice, where virtual properties are incorrectly assigned to a material being [marxCapitalCritiquePolitical2015;lukacsReificationConsciousnessProletariat2017; 133, 62, 72, 2], and the constituting role of the body in forming social cognition is shown across a variety of disciplines [37, 44, Malafouris [92]]. Yet despite the inherent abstraction of the virtual medium, little attention has been given to robots; telepresence robots exist in the physical world but are operated through a virtual medium. They provide corporeal depth to mediated contact, creating a bridge between online communication and a FtF meeting.

Moreover, the promise of social media networks to

ORCID(s):

bridge between groups and create opportunities for contact has been negated by the presence of echo chambers and filter bubbles. User of social media are typically exposed only like-minded people, and consume biased news that contribute to group polarization and an increase of prejudice [36]. Telepresence robots can occupy public spaces, transcending both physical borders and virtual bubbles, making them highly suitable for contact. We begin by reviewing the associated terminology.

2. Terminology and scope

Originally, the term *telepresence* was used by Marvin Minsky and Patrick Gunkel to describe a vision of a futuristic economy in which people perform manual, physical labor from remote locations [101]. Although the term is also used nowadays to describe a human's presence in a virtual environment [138], telepresence originally refers to the experience of being in a remote environment that is *real*, and mediated by a physical sensing agent, i.e., a *telerobot*. [25, 74]. In phenomenological terms, the experience of operating a telerobot is dubbed *re-embodiment* [38]. Today's telerobots go beyond industrial use and are deployed in social care [100], education [142], and interpersonal communication [113], utilizing the internet as the broad medium for tele-operation.

When a telerobot poses as a remote manifestation of a human operator, it is referred to as its *avatar*; the human operator could then be referred to as the *inhabiter* of that avatar. An avatar is an antonym for *agent*, a computer-controlled entity that acts autonomously without any human intervention. A telerobot is usually, however, *semi-autonomous*, combining some computer-controlled behavior with human teleoperation. A semi-autonomous telerobot is sometimes referred to as a *surrogate* [66, 106], a combination of agent and avatar.

In this article with are limiting our scope to considerations of intergroup contact when at least one of the group members is represented by a robotic avatar (or surrogate), and is physically interacting with another group member. This excludes contexts in which the robot acts a mediator [60, 129], or a simulated agent [53], or when the telerobots are interacting with each other and not with humans.

3. A path model for telepresence contact

Previous research on intergroup contact provides us with conceptual and computational tools which we can use to model the path from initial contact to the eventual reduction of prejudice toward the generalized out-group. The stages model suggested by Pettigrew [119] outlines the path to prejudice reduction: contact initially *decategories* the individual out-group member from its group, then reduces prejudice from the general out-group, and finally dissolves the border between

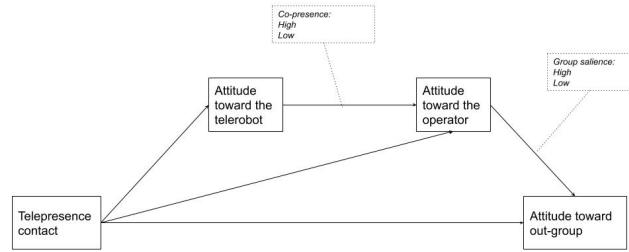


Figure 1: Telepresence Contact: Conceptual path model

in-group and out-group. Researches have also shown empirically verifiable models for various mediators and moderators involved in prejudice reduction [145, 18, 114].

Inspired by those findings, we suggest a conceptual model for telepresence contact. Due to the presence of the telerobot, we add another stage on the path to prejudice reduction fig. 1. We hypothesize that an in-group member first develops an attitude toward the robot itself before projecting it on the operating out-group member. Attitude toward the robot could be influenced by previous bias on robots or by characteristics of the particular robot which we will analyze in this article. We estimate that the perception of the robot as a representation of the operator's agency is moderated by the degree of *co-presence*. Initially formulated by Goffman as a measure of our awareness to another human being in our physical space [46], the term is now used in literature to measure the feeling of "togetherness" in mediated communication, virtual [135, 27, 12], and physical [67, 29]. Co-presence differs from the term *social presence* insofar as social presence refers to people's perception of the medium as a social sphere, rather than their recognition of sharing a space with another [21, 111]. Finally, as previous research on intergroup contact suggests, a generalized attitude toward the out-group will be moderated by the level of group salience apparent in the conversation.

4. Telepresence systems

A communication event that is mediated by telepresence robots could manifest in one of many different architectures. A single architecture could be defined as a *telepresence system*.

Utilizing concepts from Paynter's generalized systems theory [115, 50], two types of interaction that occur in telepresence, *signal* and *physical*. *Physical* interactions refer to real word interactions between elements sharing a physical environment, such as a hand-shake or holding an object. *Signal* interactions, on the other hand, occur on an abstract level. They represent a uni-

directional logical flow of cause and effect; for example, text that is typed on one end of an online communication and appears on the other end.

Based on these concepts, we identify three different types of telepresence systems see 2:

1. *Asymmetric*: The most commonly used system for telepresence communication. One participant is represented by a telerobot and is operating it from a remote location, typically their home, using a computer or mobile device. The other participant is interacting with the robot in a shared physical space. The system is asymmetric because the interaction experience is qualitatively different for the operator and the interlocutor. Common implementations include industry robots, surgical robots, office-based telepresence and social service robots.
2. *Symmetric bidirectional*: In this system both participants are operating a telerobot, which serves both as an avatar for their conversation partner as well a control interface for their telerobot. The interface, in that case, is completely transparent to the operator and typically uses body tracking. No control feedback is available to the operator and operation may commence without any knowledge of the remote telerobot by the participants. This type of system is more resource-heavy than the asymmetric type, and currently only a few implementations exist as prototypes and proofs-of-concept [106]. The lack of feedback in this system also heavily restricts the ability of the operator to travel around the space, as they are unable to see the remote environment.
3. *Symmetric unidirectional*: In this system both participants are operating a telerobot via a control interface, without physically interacting with an avatar. The two robots share the same space, while the participants are separated. Common implementations of this system include cooperative multi-robot tasks [134] and *Robot combat* competitions [30]. As in the asymmetric system, operation is usually performed from a computer or mobile device.

While the *symmetric unidirectional* system may facilitate contact in cooperative or confrontational events in front of an audience, since the scope of this article involves at least one person physically interacting with a robot, we would turn the focus toward the first two systems.

4.1 The handshake paradigm

To further illustrate the mechanics of the two systems, we will use the common example of the handshake. The handshake is an important and common social gesture in most parts of the world [126]. It is a reciprocal action, beginning with one person reaching out in an

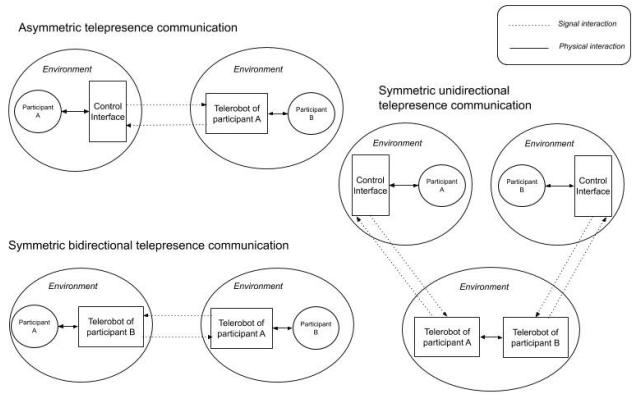


Figure 2: Systems of telepresence communication

open request and then reciprocated by the partner. It is also a haptic gesture in which subtle forms of touch can have a great social significance [58].

In an asymmetric system, the telerobot is usually equipped with a camera that streams the robot's vision back to the operator. When the interaction partner reaches out their hand, the operator will see this via the interface. The operator uses the control interface to reciprocate and have the robot reach out its hand. Movement is initiated with the push of a button in the most basic interface, or by moving a motion tracking device in a more advanced one. Once the interaction partner touches the robot, the handshake is picked up via a pressure sensor on the robot's palm and is transmitted back to the control interface. The touch could manifest as a screen flicker or a sound effect on the operator's computer, or as a vibration in a motion tracking device.

In a symmetric system, both of the participants interact only with the telerobot in front of them. The handshake is initiated when one participant reaches their hand toward the robot. The gesture is picked up by a body tracker and causes the opposing robot to mirror the behavior of its operator and reach out its hand to the interaction partner on the other side. The partner reciprocates and now both robots have their hands reached out. When touch sensors on the robots detect that both participants are now shaking their respective robot's hand, a squeezing actuation commences in both robots to generate the feeling of the handshake for both partners.

4.2 A CMC component in asymmetric systems

Whether participant is interacting with a telerobot, or with a control interface in a symmetric or asymmetric system, telepresence communication is always mediated by a computer (CMC: *Computer Mediated Communication*). Interactions can be therefore anonymous, visibility and identity can be manipulated and

conversation content can be augmented. This implies that some social psychology models that apply to CMC, would apply to one side of an asymmetric telepresence system, the side that is operating the robot via a control interface. The participants that are interacting with the robot (both sides in case of a symmetric system), are experiencing an interaction that resembles more of an FtF encounter than a mediated one. Insofar as the robot is perceived as an avatar rather than a medium, i.e a high co-presence, we hypothesize that the experience would drift further away from CMC and closer to FtF.

A number of CMC models that affect the outcome of contact and the reduction of prejudice in conversation have been identified by researchers[147]. The *deindividuation* model warns that anonymity may release a person from social regulation and norms, leading to a negative effect on the conversation. The SIDE theory [136] provides a contrasting view in which a depersonalized encounter increases group salience, as it motivates individuals to act under a group context while pronouncing enhanced norms and tropes. Models such as SIP and *hyperpersonal* communication [146] advocate that more intimate interpersonal relations may form in online contact because of the need for the participants to make up for the lack of non-verbal cues. Having a strong CMC component only on one end of an asymmetric system could create a mix of behaviors in which one side experiences stronger senses of *agency* (the sense that I am the initiator of an act) and *ownership* (the sense that it is my body that is moving) in the conversation than the other [43, 33]. Ideally, this could evolve into a positive outcome in which an out-group member on the side of the operator enjoys the benefits of anonymity and distance, while the in-group member is fully engaged in physical contact, assimilating new meanings and dissolves existing bias. Realistically however, the asymmetric nature of the conversation alone can have implications on contact, as we hypothesize in the following section.

4.3 Equality in telepresence systems

One of Allport's basic conditions for positive intergroup contact is having an equal status between group members, such as an equal position in a workplace context [3, p.274]. The condition can be further differentiated to equality in status or in *situation*, i.e in contact only; a more realistic condition when group members carry persistent power differences and are in long-lasting and deep-rooted conflicts [94].

It was shown that having symmetry and equality in communication, such as equal turn-taking and mutual respect, is beneficial for contact in groups that are in asymmetric conflict such as Israelis and Palestinians [93]. However, this does not mean that existing power relations should be ignored during contact. In fact, it

was found that a communication model that "emphasizes the conflict and power relations between the sides", although more risky, produces a higher equality within the contact [94]. Furthermore, equality emerges when the interaction focuses on fulfilling the different needs of the group members. While an advantaged group member needs to feel morally acceptance, a disadvantaged group requires a feeling of empowerment [130]. A failure to address inequality in status during contact may even hinder ensuing motivation for social change [125].

Symmetric telepresence systems provide the hardware foundation for equal contact. As will be shown in the proceeding sections, it is up to the software and organizational implementation to make sure that the conversation is handled in an equal manner. An asymmetric system, however, tilts the balance toward one end, and may have implications for contact.

In a plausible scenario, a disadvantaged-group member is operating a telerobot anonymously from their home, while the advantaged-group member is interacting with it in a public space. The operator may feel free of social anxiety in their comfort zone, and empowered by the ability to see through the robot's camera and not be seen. This in turn may encourage bringing up more difficult topics related to conflict during the conversation. However, the asymmetry in the participant's visibility may also bring up anxiety and discomfort. In our initial test for intergroup telepresence contact, participants of both sides expressed discomfort regarding the unequal asymmetric setting. One member of a minority group noted that they felt as if they are a government official investigating their exposed partners [117, p.132].

To summarize, the easier to implement asymmetric telepresence system may create both negative and positive perturbances in contact, and therefore should be designed with great consideration. While a symmetric system provides the foundation for equal grounds, leaving it up to further design choices to maintain this balance.

4. The use of a 2D display and the dual-ecologies problem

The telepresence robot market is rapidly growing, and is predicted to accelerate even more in the upcoming years with increase needs for remote working and social services ¹. Telerobot forms are continuously branching into new directions, but as of 2020 the dominant form remains to be that of a tablet device attached to a motoric vehicle [83] (See fig. ??). The tablet typically displays the operators head, as in a Skype call. Examples from market leaders include *Double Robotics* ²,

¹<https://www.marketwatch.com/press-release/telepresence-robots-market-size-2020-to-showing-impressive-growth-by-2024-industry-trends-share-size-top-key-players-analysis-and-forecast-research-2020-04-20>

²<https://www.doublerobotics.com/>

*Mantaro*³ and Revolve Robotics⁴. Such telepresence robots are oriented for remote offices and public service environments such as hospitals or schools.



width=30%}

In a contact situation, it is likely that the telerobot operator would want to remain anonymous, or at least present themselves differently than their actual form. But there is a more fundamental, ontological issue with the use of a 2D display on a telepresence robot. The problem was described as the “dual ecologies” problem, and was investigated in a case-study by Choi and Kwak [28]. In their study, a remote partner’s presence in a tablet-based video call was found to be stronger when it is disembodied (tablet only) than when it was attached to wheeled robotic body. The authors explain this by referring to the different ecologies present on the same robot: One is a 2D projection of the operator’s remote location and another is the physical presence of the robotic body in a shared space. They suggest that the receiver of communication experiences confusion, having to interact simultaneously with the immediate environment and with the depiction of the remote environment.

Our initial test also show that even a utilitarian use of a display on the telerobot’s body to perform a task, was disruptive to the participants’ sense of co-presence. Instead, participants reported reverting to the

experience of using a phone-like device while they were interacting with the display, despite having to touch the robot to initiate actions in the virtual interface [117, p.127]. Back projection solutions such as those of *Furhat Robotics*⁵ attempt to solve this incongruence by projecting 3d-mapped virtual information directly on the robot’s surface. While this may moderate the negative effects of the display, it does not entirely address the dual ecologies problem of having two different spatial sources combined into one.

Therefore, the use of a display should be planned carefully for intergroup contact. Alternatively, the robot’s appearance could be designed with an emphasis on uniformity and consistency. Considerations for appearance are discussed in the following sections.

5. Telerobot design for contact

5.1 Visual appearance

The affect of a robot’s appearance on a human’s attitude toward it has been studied extensively in literature, predominantly in studies of Human-Robot Interaction (HRI) and social robotics [49]. A pivotal discussion revolves around the question of *anthropomorphism*: the degree in which a robot’s appearance and behavior resembles that of a human. Current literature paints a picture that is manifold [40]; while anthropomorphic features may increase a human’s empathy and acceptance of the robot, the effect is context-dependent, culture-dependent, and works only to some extent. In some cases people have preferred pet-shaped over human-like robots, and in other cases a human shape generated a negative attitude toward the robot. One over-reaching theory on robot-human-likeness has been extensively covered in literature: Mori’s theory of the *Uncanny valley* [104]. The theory suggests that a human’s affinity toward a robot grows as it displays more human-like features, but at some point falls into a valley of uncanniness when the robot is *too human* yet noticeably *inhuman*. The affinity rises back up only when theoretically no difference between the robot and a human can be discerned.

The uncanny, as it was originally described by Freud [42], invokes emotions as *morbid anxiety*; anxiety being a known negative mediator for the outcome of contact [120, p.767]. However, the context of telepresence adds another dimension to this question. It is not a given that an attitude formed toward the telerobot would naturally project onto the operator. We therefore suggest to extend the model of the uncanny to support telepresence by adding another dimension to the equation. As noted, the level of Avatar co-presence moderates the transference of affect from the avatar to the inhabitant. Adding another multiplier to the uncanny valley equation would then reflect the estimated positive effect of avatar co-presence (see fig. 3).

³<http://www.mantarobot.com/products/teleme-2/index.htm>

⁴<https://telepresencerobots.com/robots/kubi>

⁵<https://furhatrobotics.com/>

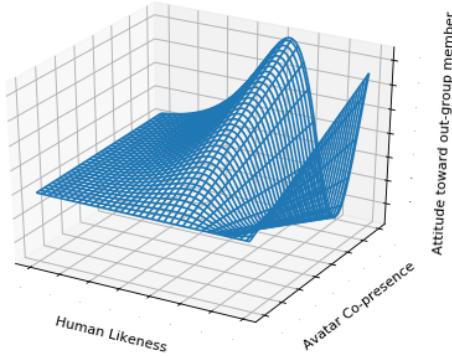


Figure 3: Extended uncanny valley model for intergroup contact

Problematizing appearance even further, findings indicate that the level co-presence is also moderated by the level of anthropomorphism. Research by Groom et al [?] suggests that robot operators have a greater sense of *self extension* on their avatar when it is nonhuman. They explain this finding by the fact that a humanoid robot assumes its own identity, making it harder for the operators to project themselves on the robot. But on the other hand of the conversation, research by Kuwamura et al [84] found that the interlocutors experience a distortion of their partner's personality when it is conveyed through a nonhuman. Compared to a humanoid telerobot, participants interacting with a stuffed-bear felt confused, and had difficulty imagining that they are talking to a human.

Weighing in the above factors, we hypothesize that a non-anthropomorphic approach has more benefits than the anthropomorphic approach. A combination may also work well, maintaining some expressive anthropomorphic features. We suggest to circumvent the problem of consistency between the telerobot's appearance and the human operator by adding outside cues and explanations that remind the interlocutor of the human inhabiter. We will expand it more in the section regarding group salience. When using a non-anthropomorphic appearance, such as zoomorphic or caricatured, in intergroup contact, caution should be maintained to not demean the participants or portray in-equality. As noted by Lee et al[87], features such as the robot's height in respect to a human may influence the perception of the human behind the robot.

5.4 Voice

Research on user attitude toward an avatar's voice shows that as with visual appearance, one must strike a balance between relatedness and consistency. Lee and Nass studied the sense of social presence of avatars with machine-generated voices [88] in respect to their



Figure 4: From left to right: Telenoid, Elfoid, Retreived from <http://www.geminoid.jp/en/robots.html>



Figure 5: Geminoid and Professor Hisroshi Ishiguro. Retreived from <http://www.geminoid.jp/projects/kibans/resources.html>

personality (introvert or extrovert). When an avatar's voice personality is closer to that of the interlocutor, the sense of social presence is increased. However, the consistency of the voice with the portrayed personality of the avatar is essential. Social presence drops if the voice is incongruent with the textual personality. Another study by Mitchel et al [102] found that a mismatch between the voice and the face of a talking head generated an uncanny sensation. A human with a synthetic voice felt as uncanny as a robot with a human voice.

In a telepresence robot contact, an operator may choose to use their own voice or a synthetic voice that does not disclose their personality, gender and culture. They may also use a synthetically cloned voice that is highly similar to their natural voice [69]. A synthetic voice also adds modalities for speech augmentation and language translation as we discuss in the section regarding interaction modalities. In order to avoid anxiety and confusion, we hypothesize that an optimal voice would be one that takes the telepresent nature into account, reminding the interlocutor that there is a human operator inhabiting the avatar, but also does not diverge abruptly from the physical form of the avatar.

5.2 Designing with group salience in mind

Naturally, the process toward forming a generalized opinion toward the out-group does not end at the interpersonal level. The most widely agreed-upon moderator for generalizing a positive attitude to the intergroup level is *group salience*, the degree in which the parti-

pants' group identity is salient during contact. Earlier research on contact found a bipolar, antagonistic correlation between group identity salience and interpersonal relationship [19]: When the group identity is so present, it is harder for the participants to find their commonalities and create new bonds and friendships. Later research, however, suggests that there is now a sufficient evidence that group and interpersonal salience can remain high simultaneously [18]. One approach, suggested by Pettigrew [119], is to expose group identities gradually, starting with a low salience, allowing initial contact to form, and increasing it over time.

Group identity can be transmitted through a variety of channels in telepresence, beginning with the design of the avatar; its appearance, voice, and its surroundings, and proceeding into the content of the interaction. Although the physical medium is not as easily manipulated as a virtual environment [52], there is still room for creativity in the way group identity is revealed during contact. For example, an avatar may have a non-humanoid appearance, but still maintain group identity through group symbols, cues, language, and so forth. It may speak in an accent, wear typical accessories or flaunt national colors. The freedom to use material objects brings up new design possibilities that are not available in an online encounter. Group cues could be positioned in subtle ways, so that they are gradually revealed by the interlocutor. If the initial appearance and behavior of the robot is engaging enough, an interpersonal bond may be formed despite the presence of group-related cues.

5.3 Materiality

An important factor in robot design is the choice of materials. In industrial robots, the material would be chosen *functionally*, in accordance with the task at hand. In robots designated for human interaction, instead of materials, we examine *materiality*, conceptualized by N. Katherine Hayle as "physical qualities that present themselves to us" [55]. Materiality comes into play in two main aspects of a robot's constitution: 1) The outer skin: the part of the robot that touches and is being touched, and 2) Actuation: the material that actuates, generating the robot's movements. With the former, we can place materials on a scale of firmness and rigidness; how soft it feels to touch the material. With the latter, we can define a scale of flexibility and linearity, describing the nature of the material's movement.

Previous research in social robotics supports the use of soft materials for the outer skin of robots, especially in interaction with children [81] and in elderly care [16, 79]. The reason being the presence of *affective touch* between the robot and humans [78, 139]. The human body and another natural forms are inherently soft, and co-exist better with other materials that are soft [35]. Nevertheless, here too, carrying an object closer to the realm of the living may invoke an uncanny feeling, for

example when touching a smooth, soft, material that is also cold [150, 109].

A soft touch on the surface doesn't necessarily imply a softness as a whole. For example, a gripping robotic hand made from powerful servo motors that are wrapped in a soft skin could still easily, and inadvertently, crush soft tissue. *Soft Robotics* is a rapidly developing research field for robots that operate on soft materials down the level of actuation [10]. Research is often, but not always, bio-inspired; commonly used materials are fabrics and silicone rubbers while the most common form of actuation is pneumatic: applying air pressure or vacuum. Presently, the largest consumers of soft robotics are the medical industry, utilizing the soft materials for invasive and surgical procedures. Nevertheless, uses and appeal of soft robots for human interaction are being studied [14, 71, 153], uniformly arriving to the conclusion that the smooth, organic-like motion of soft actuators has an expressive quality, and affects humans in a manner highly different from traditional robotics. Our small-scale test on soft robotic telepresence has reached similar conclusions [117, p.126]. As predicted by Mori, movement has an effect on the uncanny, and soft actuation has potential to fall into the a negative category. If designed correctly, movement can bestow a robot with a relatable, expressive quality, despite it having a non-anthropomorphic appearance [59].

All things considered, a soft approach deems appropriate for telepresence contact. Insofar as the presentation of the avatar remains balanced in respect to the principles outlined in this section, adding soft movement and touch may increase empathy and intimacy in contact and lead to a more positive outcome. A soft approach also presents a notion of safety, and an inability to cause harm; a desirable climate in situations of intergroup conflict.

TODO: Robot's Height, in equality?

5.5 Avatar co-design

The ability to customize an avatar is widespread in video games and social media applications. It was shown to improve a user's engagement in the platform [108], and to have consequence on a user's social behavior. According to studies of *self-extension* in the digital world [11, 99] users tend to have an increased self-confidence and behave more extroverted when they perceive their avatar as more attractive. The effects of self-extension may be so intense that the changes in behavior remain after even usage.

In robotics, customizing an avatar is more complex than in a virtual environment. Possibilities for design are constrained by the hardware platform of the robot, requiring a form of co-design between the robot engineers and the user. Assembling the robot requires physical effort and basic knowledge in mechatronics, and cannot be performed intuitively as in a video game. Nevertheless, involving users in designing and assembling their

robotic avatar may have great benefits. In the study by Groom et al[47], it was shown that operators had a greater sense of self-extension to a robot that was assembled by them, rather than by another. Robots were also successfully co-designed with children as the target users. The YOLO robot focused on creativity and storytelling, allowing children to design behaviors and movements [4], while the PAL involved children in designs for diabetes self-management [57]. Co-design methods also improved the general attitude of students toward robots in educational settings [123].

In the context of contact, and especially in situations of intergroup conflict, co-designing avatars may have even greater virtues. Participants could control their representation and its behavior, considering how they wish to be seen by the other side; thus, supporting ‘controlled’ means for escalation and de-escalation that has been found beneficial in other media-based intergroup contact projects [151]. The assembly work in itself may be therapeutic, both as tactile experience [131] and as a self-expressive arts [105]. Finally, a participatory approach for robot-building holds great potential for empowering oppressed groups and minorities by providing meaningful education in modern communication technologies.

6. Interaction modalities

As described in the beginning of this article, the term *physical interaction* delineates the missing element in a signal-based, online interaction. Physical interactions occur in a real environment where matter creates an energy bond with matter. This includes *embodied interactions*: subject-to-object interactions between a human body and an object such as a robot, and object-to-object interactions, e.g. when the robot is interacting with other objects in the environment. In telepresence contact, we want to create a real sense of intersubjectivity between human and avatar, as close as possible to situated FtF encounter. Phenomenologist Merleau-Ponty coined the term *intercorporeality* to describe the strong and co-dependent physical nature of intersubjectivity [98]. Today, researchers in the field of embodied cognition are studying the specific mechanisms behind this phenomenon.

In this section we explore possible affordances [45, p.285] in telerobot design, and theorize over different modalities [82]: modes for interaction, that may assist in intergroup contact. We detail suggestions toward implementation in the two types of systems discussed: symmetric and asymmetric. For every implementation, we consider the level of agency allotted to the operator, and conversely the level of autonomy required from the robot. While we strive to grant full agency to the operator, some modalities require an intervention from the robot to ensure a smooth experience, while other modalities may benefit from being augmented with the

robot’s capabilities.

6.1 Movement in space

The ability to move a body in space is what mainly distinguishes a robot from other interactive technologies. However, not all robots have the same degrees of freedom and granularity when it comes to movement. In the field of social robots, *mobile robots* typically travel around using wheeled motion. Examples include service robots, such as Pepper⁶ and Samsung bots⁷, and telepresence robots such as Double Robotics⁸ and Beam⁹. Other robots only move their body while remaining stationary in place. For example care robot such as PARO¹⁰, and Ishiguro laboratories’ telepresence robots¹¹. Due to the complexity of maintaining both modalities in interaction, mobile robots usually keep a physical distance from the user, interacting mostly using voice and visual information as they travel around. Stationary robots, on the other hand, tend to rely more on haptic interaction, allowing the user to hold and touch them. Only a few robots attempt to combine both modalities, such as Teo [15].

Touch-based interactions’ affective role in human-robot relations [6, 78] renders them meaningful for contact, as it was shown that the outcome of contact is tightly integrated with affective measures. Thus it may prove beneficial to place touch modalities in a higher priority than the ability to move around in space. Moreover, in the cases of symmetric telepresence, moving around is nearly impossible, as the operator does not have a dedicated interface for traveling, and movement can rely only on limited body tracking. Even in asymmetric systems, complexities added by the need to navigate a robot using a camera may divert attention from the main task at hand, having a meaningful interpersonal interaction.

One exception in which acquiring mobility may be paramount is in situations of asymmetrical conflict where the oppressed group suffers from movement restrictions in their daily life, as is the situation in Palestine [17]. In such case, an oppressed group may get empowered by having the ability to travel around, using their telepresence robot, in areas that are normally out of their reach. This, in turn, may contribute to greater sense of equality and confidence within the conversation.

6.2 Nonverbal communication and emotional expression

Nonverbal communication (NVC) plays a substantial role in our day-to-day interactions. In a pioneering work by Michael Argyle [8], nonverbal signals were listed as follows: facial expressions, gaze, gestures, posture,

⁶<https://www.softbankrobotics.com/emea/en/pepper>

⁷<https://research.samsung.com/robot>

⁸<https://www.doublerobotics.com/>

⁹<https://suitabletech.com/>

¹⁰<https://www.parorobots.com/>

¹¹<http://www.geminoid.jp/en/robots.html>

bodily contact, spatial behavior, appearance, non-verbal vocalizations and smell. A telerobot has the capacity to use all of those signals, but the main challenge lies with the reliable transmission of signals from the operator's intention to the re-embodied form of the robot. Signals have to be picked up accurately from the operator, and then translated to the telerobot's body to convey the same meaning.

Argyle categorizes types of NVCs according to their level of awareness within the sender and the receiver. The majority of NVCs are defined as *mostly unaware* on part of the sender and *mostly aware* on the receiver side. For example: we are only slightly conscious of the fact we are smiling during a conversation, but the sentiment is registered more attentively with our conversation partner.

Since NVCs communicate emotion, they should be handled with great care in telepresence contact. While unaware emotional signals can be detected on the operator using facial recognition, prosodics, and body tracking, modern deep learning systems are still subject to noise and occasional error, and can only recognize generalized basic emotions[see 152, 63]. Theories of universality in facial expression, originating in Darwin's account of emotional expression [?] suggest a common basis for expression across cultures [51], but the findings are inconclusive [68]. A mistake in an assumption of a participant's emotional state could lead to confusion and frustration in the conversation, therefore it may be safer to rely strictly on explicit gestures made by the operator with full awareness. Expression of emotions in an asymmetric system could be invoked by the operator using Emojis[75] or other dedicated buttons. In a symmetric system, explicit body gestures could be used, or touch-based interactions such as a pat or a stroke on the robot's body.

Once the operator's intention is picked up, it can be expressed by the avatar, and here too is a risk of confusion. If using a humanoid avatar, a combination of gaze and smooth eyebrow motion can display expressive facial expressions, as shown in the SEER robot [143]. , and studies of emotion expression in virtual humanoids avatar have shown that they add an element of confusion and mis-representation [116]. When using a non-humanoid, facial expressions are very limited. Instead, body movement can be used to express emotions. Laban's movement analysis [85] (LMA), a framework from the field of dance was shown to be applicable for robots [107, 95]. LMA analyzes movement components such as *effort* and *shape*, that can be assigned to basic emotions [128]. Using dance as an inspiration for robotic expression also benefits from its cross-culture universality [132]. Another model, the PAD emotional state model [96], was successfully applied to robot movement design [7, 110]. PAD describes emotions in three dimensions: Pleasure, Arousal and Dominance. Defining and evaluating the robot's emotional state on the PAD

scale allows designers to correlate emotions with specific mechanical movements.

Some signals, such as shifts in gaze and body orientation, are completely unaware during conversation, yet they have an impact on the turn-taking and attention signaling [77, 124]. Smooth turn-taking promotes the sense of equality in contact, and was demonstrated to benefit human-robot interaction [144, 86]. Insofar as those actions cannot be explicitly pronounced by the operator without impairing the flow of the conversation, they need to be a part of the robot's semi-autonomous functionality. When physically interacting with a robot, end-of-turn can be predicted using a number of tracking modalities [1]. In an asymmetric interface when communication is simply typed or clicked, end-of-turn can be inferred directly from the operator, and signals such as a typing indicator could be sent. In the opposing end however, the camera-based view does not properly transmit turn-taking signals, and may break the flow of the conversation, as is widely seen in video conferencing applications[41].

6.3 Verbal communication

Bridging conflict narratives and reducing prejudice between groups seems virtually impossible without the use of language. Whether one holds a more embodied approach of speech [97], or views language as the logical foundation of the lifeworld [48], there is a consensus on its constitutive role in forming intersubjectivity [56]. In fact, it was shown that language can make up for the lack of embodied interaction, and is often used creatively in CMC to overcome missing nonverbal modalities [76]. For the above reasons we consider language modalities to be essential in telepresence contact.

The outstanding benefit of mediated verbal interaction for intergroup contact is the ability to translate between different languages and dialects. This has significance since often groups in conflict do not speak a common language and communicate only through mass-media which propagates biased views. Machine translation, however, may also be destructive to cultural and political nuances [89, 34], and contemporary deep-learning translators exhibit stylistic and gender bias based on their training datasets [65, 137]. Models such as Timo Honkella's "peace machine" [61, 80] attempt to resolve this problem by preserving cultural-dependent meanings within translation. While no data exists so far on the negative implications of machine translation in intergroup contact, it is known that mediating human translators and interpreters suffer from a lack of trust by the participants who worry that their voice will not be heard [103, p.22]. In our initial test for automatic language translation in contact, participants have raised concerns of being misrepresented by the machine [117, p.130].

While caveats of machine translation cannot be disposed of, there are some mitigating steps that can be

taken to improve the experience of the participants. First, when using speech recognition, feedback of the recognition result in the operator's native language should be given, perhaps even allowing, at the cost of maintaining a conversation flow, to confirm the result before sending it out for translation. Second, when possible, the interface should display the confidence level of the translation before it gets sent to other side. Finally, in case a mistake was realized by the operator only after submitting, there can be a quick "oops" button that has the robot express an apologetic gesture. While mediated translation requires the use of a machine generated artificial voice that can speak the desired language, our initial tests show that participants enjoy their newly acquired ability to speak another language, and are excited to try it [117, p.130]. If used according to the above mentioned principles of design and interaction, real time language translation could be an important facilitator for telerobotics intergroup contact.

6.4 Synchrony, reciprocity and feedback

Whether communication is verbal or not, it is shown that synchrony and reciprocity facilitate interpersonal sympathy and empathy, also referred to as "social entrainment" [13, 22, 122], Stupacher et al. [140]]. Such interactions include: rhythmic movements (e.g clapping, jumping), a smooth conversation beat, synchronized dance, give-and-receive interactions, gaze synchrony, affective matching, and mimicry. The same positive effects are observed in human-robot interaction, particularly in care and therapy robots for children and the elderly [9, 91].

The biggest problem in achieving interpersonal synchrony over mediated communication is that of *latency* [26]. The unavoidable time delay due to physical distance between the participants can incite confusion and frustration when performing rhythmic and simultaneous tasks, even if the internet round-trip amounts to a few hundred milliseconds. Research in the field of online music performance is at the forefront of tackling such issues [112]. Semi-autonomous methods in the likes of action prediction, lag compensation and global metronomes enable musicians to collaborate in jam sessions from different locations around the globe. The same methods can be used to synchronize robot-mediated activities such as gesture mimicry, dance, music and cooperative gaming. Synchronize actions can be performed both in symmetric and asymmetric systems. In a symmetric system the participants would be coordinating the same action, for example clapping, while in an asymmetric system one participant would be using a control interface, for example tapping or shaking a mobile phone, while another would perform the action in front of the telerobot.

Some reciprocal actions do not require real-time synchrony between the participants and can be implemented more easily. For example, the above mentioned hand-shake paradigm and other similar gestures such

as a "high five" could be performed in a turn-based flow. This may decrease the valence of the action, but still benefit the conversation; people tend to be rather forgiving toward the sluggishness of robots. Additionally, the semi-autonomous function of the telerobot could simulate instant feedback and nonverbal reactions such as gaze shifting in response to the interlocutor's actions.

In an asymmetric system where the robot's operator is using a control interface, having instant feedback to the control actions increases the sense of agency and ownership within the operator [38]. At the high-end of the spectrum we can find advanced control systems that mix Virtual Reality and haptic feedback such as the ones for 'Robonaut' at the Johnson space center [32]. As a bare minimum, however, an operator should have visual feedback of how the robot acts in response to control commands. This can be achieved by placing a mirror on-site, providing the operator with multiple camera angles, and practicing the use of the robot in-advance. In our initial experiment, participants have expressed concern over their inability to see the facial expressions they were invoking with the robot. The lack of feedback was especially notable when performing a reciprocal action such as a hand-shake. When participants were unable to see their avatar's arms they could only guess what is happening. It is thus critical to supply ample feedback to the operator in an asymmetric system.

7. Peacebuilding scenarios

7.1 Public space interventions

Robotic avatars can be used as an advanced communication tool in organized intergroup encounters where participants are unable to meet face-to-face, but the physical presence of telerobots makes them exceptionally suitable for random public space interventions. Robots can transcend national borders and occupy spaces, with potential to reach crowds that wouldn't normally engage in intergroup contact. In a symmetric system, robots could be placed in public urban spaces of both sides and wait of passersby to initiate contact. In an asymmetric system, one robot would be placed in a public space while operators would inhabit the avatar from a remote location.

In such a scenario, one might consider whether the group identity of the telerobot's operator should be widely exposed to passersby, allowing them to make a voluntary decision to approach, or whether they would first approach the robot and only realize its group identity during conversation. A meta-analysis by Pettigrew et al [121] concluded that in contacts that had a negative outcome, it was far worse when the contact was initiated involuntarily. However, when the group identity is known, members may avoid interaction [148], even in organized interventions between groups [94]. Once again a balanced approach may work well here. Some external cues could be exposed, providing only hints about

the telerobot's identity, allowing passersby to approach voluntarily while considering the robot's group identity. A robot that pops up in the middle of a public space might be intriguing enough for some people to approach. Depending on the robot's design and on a person's bias toward robots, some may also deter from approaching this unfamiliar entity. One method to get public interest would be to equip the robot with some actions that are designed to draw a crowd, and allow operators to trigger them. For example playing a sound, a musical theme, or performing some inviting gesture. In an asymmetric system, an operator has more control on the robot's interaction with the environment. They make look around by moving a camera, or even drive around using wheeled motion. When planning a public space intervention, the designated site and its demographics should be considered. A design that is tailored to the site and context would have a higher chance to succeed.

7.2 Interaction content

Once contact if formed, the conversation between the participants can flow freely, or may be guided by the telerobot's semi-autonomous functionality. Modern types of interaction may include cooperative games and simulations that engage the participant toward a common goal, in-line with Allport's principles. Games have already shown potential in face-to-face meetings [20] and in online contact [52]. When interacting with a robot, its body parts may be appropriated as game-controllers. For example, squeezing the robot's arms, thus indirectly forming touch interactions between the participants. In games that offer visual content, it is necessary to embed a 2D display in the telerobot or place one beside it. In that case, attention of the users should be managed carefully, and content should be distinct from the avatar's body to avoid the dual ecologies problem and a reduced co-presence.

More traditional forms of conflict resolution and peacebuilding can also be supported by the robot's functionality. Seen as a mediator that supports interaction, a transformative approach to conflict mediation which emphasizes recognition and empowerment rather than problem solving [24] seems appropriate. The robot's interface may encourage the participants to express themselves freely and make sure that everyone has a right to be heard. The interaction could also make use of storytelling as a tool for the participants to disclose their narrative [31] to one another. At the forefront of technological conflict mediation, machine learning is being sought as a tool for peacebuilding that can understand complex sentiments and situations [61], predict conflict escalation before it occurs [118], and offer help to resolve issues [141]. As noted earlier, machine learning tools should always be used with caution and awareness of their training environment.

The content suggested above only scratches the sur-

face of conflict resolution and peacebuilding methods, as the possibilities for interaction are boundless. It is important to note, however, that in a telepresence context the sense of co-presence may be affected as the mediating autonomy of the robot increases. The robot may be perceived more as a middle-agent rather than an avatar, which in turn may reduce the effect of positive outcome toward its operator. Mediating functions should be kept as transparent and seemless as possible.

9. Ethical considerations

10. Summary table

References

- [1] , . Multimodal end-of-turn prediction in multi-party meetings | Proceedings of the 2009 international conference on Multimodal interfaces. <https://dl.acm.org/doi/abs/10.1145/1647314.1647332>.
- [2] Ahmed, S., 2000. Strange Encounters : Embodied Others in Post-Coloniality. Routledge, London.
- [3] Allport, G.W., 1954. The Nature of Prejudice. Addison-Wesley, Oxford, England.
- [4] Alves-Oliveira, P., Arriaga, P., Paiva, A., Hoffman, G., 2017. YOLO, a Robot for Creativity: A Co-Design Study with Children, in: Proceedings of the 2017 Conference on Interaction Design and Children, ACM, Stanford California USA. pp. 423–429. doi:10.1145/3078072.3084304.
- [5] Amichai-Hamburger, Y., Hasler, B.S., Shani-Sherman, T., 2015. Structured and unstructured intergroup contact in the digital age. Computers in Human Behavior 52, 515–522. doi:10.1016/j.chb.2015.02.022.
- [6] Andreasson, R., Alenljung, B., Billing, E., Lowe, R., 2018. Affective Touch in Human–Robot Interaction: Conveying Emotion to the Nao Robot. International Journal of Social Robotics 10, 473–491. doi:10.1007/s12369-017-0446-3.
- [7] Ardila, L.R., Coronado, E., Hendra, H., Phan, J., Zainalkefl, Z., Venture, G., 2019. Adaptive Fuzzy and Predictive Controllers for Expressive Robot Arm Movement during Human and Environment Interaction. International Journal of Mechanical Engineering and Robotics Research , 207–219doi:10.18178/ijmerr.8.2.207–219.
- [8] Argyle, M., 2013. Bodily Communication. Routledge.
- [9] Aucouturier, J.J., Ikeuchi, K., Hirukawa, H., Nakaoaka, S., Shiratori, T., Kudoh, S., Kanehiro, F., Ogata, T., Kozima, H., Okuno, H.G., Michalowski, M.P., Ogai, Y., Ikegami, T., Kosuge, K., Takeda, T., Hirata, Y., 2008. Cheek to Chip: Dancing Robots and AI's Future. IEEE Intelligent Systems 23, 74–84. doi:10.1109/MIS.2008.22.
- [10] Bao, G., Fang, H., Chen, L., Wan, Y., Xu, F., Yang, Q., Zhang, L., 2018. Soft robotics: Academic insights and perspectives through bibliometric analysis. Soft robotics 5, 229–241.
- [11] Belk, R.W., 2013. Extended Self in a Digital World: Table 1. Journal of Consumer Research 40, 477–500. doi:10.1086/671052.
- [12] Bente, G., Rüggenberg, S., Krämer, N.C., Eschenburg, F., 2008. Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations. Human Communication Research 34, 287–318. doi:10.1111/j.1468-2958.2008.00322.x.
- [13] Bernieri, F.J., Rosenthal, R., 1991. Interpersonal coordination: Behavior matching and interactional synchrony.
- [14] Bewley, H., Boer, L., 2018. *Designing Blo-nut*: Design Principles, Choreography and Otherness in an Expressive

- Social Robot, in: Proceedings of the 2018 on Designing Interactive Systems Conference 2018 - DIS '18, ACM Press, Hong Kong, China. pp. 1069–1080. doi:10.1145/3196709.3196817.
- [15] Bonarini, A., Garzotto, F., Gelsomini, M., Romero, M., Clasadonte, F., Yilmaz, A.N.C., 2016. A huggable, mobile robot for developmental disorder interventions in a multi-modal interaction space, in: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 823–830. doi:10.1109/ROMAN.2016.7745214.
- [16] Broekens, J., Heerink, M., Rosendal, H., 2009. Assistive social robots in elderly care: A review. *Gerontechnology* 8, 94–103. doi:10.4017/gt.2009.08.02.002.00.
- [17] Brown, A.P., 2004. The Immobile Mass: Movement Restrictions in the West Bank. *Social & Legal Studies* 13, 501–521. doi:10.1177/0964663904047331.
- [18] Brown, R., Hewstone, M., 2005. An integrative theory of intergroup contact. *Advances in experimental social psychology* 37, 255–343.
- [19] Brown, R.J., Turner, J.C., 1981. Interpersonal and intergroup behaviour. *Intergroup behavior*, 33–65.
- [20] Brynen, R., Milante, G., 2013. Peacebuilding With Games and Simulations. *Simulation & Gaming* 44, 27–35. doi:10.1177/1046878112455485.
- [21] Bulu, S.T., 2012. Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education* 58, 154–161. doi:10.1016/j.compedu.2011.08.024.
- [22] Burgoon, J.K., Dillman, L., Stem, L.A., 1993. Adaptation in Dyadic Interaction: Defining and Operationalizing Patterns of Reciprocity and Compensation. *Communication Theory* 3, 295–316. doi:10.1111/j.1468-2885.1993.tb00076.x.
- [23] Burgoon, J.K., Hoobler, G.D., 1994. Nonverbal signals. *Handbook of interpersonal communication* 2, 229–285.
- [24] Bush, R.A.B., Folger, J.P., 2004. The Promise of Mediation: The Transformative Approach to Conflict. John Wiley & Sons.
- [25] Campanella, T., 2000. Eden by wire: Webcameras and the telepresent landscape, pp. 22–46.
- [26] Campbell, J., 2015. Interpersonal Coordination in Computer-Mediated Communication, in: Encyclopedia of Information Science and Technology, Third Edition. IGI Global, pp. 2079–2088.
- [27] Casanueva, J., Blake, E., 2001. The effects of avatars on co-presence in a collaborative virtual environment .
- [28] Choi, J.J., Kwak, S.S., 2016. Can you feel me?: How embodiment levels of telepresence systems affect presence, in: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), IEEE, New York, NY, USA. pp. 606–611. doi:10.1109/ROMAN.2016.7745180.
- [29] Choi, J.J., Kwak, S.S., 2017. Who is this?: Identity and presence in robot-mediated communication. *Cognitive Systems Research* 43, 174–189. doi:10.1016/j.cogsys.2016.07.006.
- [30] Clarkson, M., Foreword By-Dwyer, B., 2002. BattleBots: The Official Guide. McGraw-Hill, Inc.
- [31] Coburn, C., 2011. Storytelling as a Peacebuilding Method, in: Christie, D.J. (Ed.), *The Encyclopedia of Peace Psychology*. Blackwell Publishing Ltd, Oxford, UK, p. wbepp268. doi:10.1002/9780470672532.wbepp268.
- [32] Cola, C., . Small World Machines.
- [33] Cole, J., Sacks, O., Waterman, I., 2000. On the immunity principle: A view from a robot. *Trends in Cognitive Sciences* 4, 167. doi:10.1016/S1364-6613(00)01459-5.
- [34] Cronin, M., 2012. Translation in the Digital Age. First ed., Routledge. doi:10.4324/9780203073599.
- [35] Danese, E., 2003. Soft machine. *Machines That Become Us: The Social Context of Personal Communication Technology*, 267–276.
- [36] Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., Quattrociocchi, W., 2016. Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports* 6, 37825.
- [37] Dewey, J., 1986. Experience and education, in: *The Educational Forum*, Taylor & Francis. pp. 241–252.
- [38] Dolezal, L., 2009. The Remote Body: The Phenomenology of Telepresence and Re-Embodiment. *Human Technology* 5, 208–226.
- [39] Eckert, C., Isaksson, O., Hallstedt, S., Malmqvist, J., Rönnbäck, A.Ö., Panarotto, M., 2019. Industry Trends to 2040, in: *Proceedings of the Design Society: International Conference on Engineering Design*, Cambridge University Press. pp. 2121–2128.
- [40] Fink, J., 2012. Anthropomorphism and human likeness in the design of robots and human-robot interaction, in: *International Conference on Social Robotics*, Springer. pp. 199–208.
- [41] Fischer, K., Tenbrink, T., . Video conferencing in a transregional research cooperation: Turn-taking in a new medium , 17.
- [42] Freud, S., 1919. The Uncanny. *Imago* 1, 1–21.
- [43] Gallagher, S., 2000. Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences* 4, 14–21.
- [44] Gallagher, S., 2006. How the Body Shapes the Mind. Clarendon Press.
- [45] Gibson, J.J., 1966. The senses considered as perceptual systems. .
- [46] Goffman, E., 2008. Behavior in Public Places. Simon and Schuster.
- [47] Groom, V., Takayama, L., Ochi, P., Nass, C., 2009. I am my robot: The impact of robot-building and robot form on operators, in: *Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction - HRI '09*, ACM Press, La Jolla, California, USA. p. 31. doi:10.1145/1514095.1514104.
- [48] Habermas, J., 1984. *The Theory of Communicative Action*. volume 1. Beacon Press, Boston.
- [49] Hancock, P.A., Billings, D.R., Schaefer, K.E., Chen, J.Y.C., de Visser, E.J., Parasuraman, R., 2011. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53, 517–527. doi:10.1177/0018720811417254.
- [50] Hannaford, B., 2000. *Feeling Is Believing: A History of Telerobotics. The Robot in the Garden: Telerobotics and Telepistemology in the Age of the Internet*. Edited by Ken Goldberg. The MIT Press.
- [51] Hareli, S., Hess, U., 2017. Facial Expressions and Emotion, in: Zeigler-Hill, V., Shackelford, T.K. (Eds.), *Encyclopedia of Personality and Individual Differences*. Springer International Publishing, Cham, pp. 1–7. doi:10.1007/978-3-319-28099-8_518-1.
- [52] Hasler, B.S., Amichai-Hamburger, Y., 2013. Online Intergroup Contact, in: Amichai-Hamburger, Y. (Ed.), *The Social Net*. Oxford University Press, pp. 220–252. doi:10.1093/acprof:oso/9780199639540.003.0012.
- [53] Hasler, B.S., Hirschberger, G., Shani-Sherman, T., Friedman, D.A., 2014. Virtual Peacemakers: Mimicry Increases Empathy in Simulated Contact with Virtual Outgroup Members. *Cyberpsychology, Behavior, and Social Networking* 17, 766–771. doi:10.1089/cyber.2014.0213.
- [54] Hasson, Y., Schori-Eyal, N., Landau, D., Hasler, B.S., Levy, J., Friedman, D., Halperin, E., 2019. The enemy's gaze: Immersive virtual environments enhance peace promoting

- attitudes and emotions in violent intergroup conflicts. PLOS ONE 14, e0222342. doi:10.1371/journal.pone.0222342.
- [55] Hayles, N.K., 2014. Speculative aesthetics and object-oriented inquiry (OOI). *Speculations: A journal of speculative realism* 5, 158–79.
- [56] Haysom, K., 2009. Communicating Depth: Habermas and Merleau-Ponty on Language and Praxis. *Political Theory* 37, 649–675.
- [57] Henkemans, O.A.B., Neerincx, M.A., Oleari, E., Pozzi, C., . Co-design of the PAL robot and avatar that perform joint activities with children for improved diabetes self-management , 3.
- [58] Hillewaert, S., 2016. Tactics and tactility: A sensory semiotics of handshakes in coastal Kenya. *American Anthropologist* 118, 49–66.
- [59] Hoffman, G., Ju, W., 2014. Designing Robots With Movement in Mind. *Journal of Human-Robot Interaction* 3, 89–122. doi:10.5898/JHRI.3.1.Hoffman.
- [60] Hoffman, G., Zuckerman, O., Hirschberger, G., Luria, M., Shani-Sherman, T., 2015. Design and Evaluation of a Peripheral Robotic Conversation Companion, in: 2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 3–10.
- [61] Honkela, T., 2017. Rauhankone: Tekoälytutkijan testamentti .
- [62] Horkheimer, M., 2004. Eclipse of Reason. Rev. ed. ed., Continuum, London ; New York.
- [63] Hossain, M.S., Muhammad, G., 2019. Emotion recognition using deep learning approach from audio–visual emotional big data. *Information Fusion* 49, 69–78. doi:10.1016/j.inffus.2018.09.008.
- [64] Hourcade, J.P., Bullock-Rest, N.E., 2011. HCI for peace: A call for constructive action, in: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11, ACM Press, Vancouver, BC, Canada. p. 443. doi:10.1145/1978942.1979005.
- [65] Hovy, D., Bianchi, F., Fornaciari, T., . Can You Translate that into Man? Commercial Machine Translation Systems Include Stylistic Biases , 5.
- [66] Hughes, C.E., 2014. Human Surrogates: Remote Presence for Collaboration and Education in Smart Cities, in: Proceedings of the 1st International Workshop on Emerging Multimedia Applications and Services for Smart Cities - EMASC '14, ACM Press, Orlando, Florida, USA. pp. 1–2. doi:10.1145/2661704.2661712.
- [67] Hwang, J., Sangyup Lee, Sang Chul Ahn, Hyoung-gon Kim, 2008. Augmented robot agent: Enhancing co-presence of the remote participant, in: 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE, Cambridge, UK. pp. 161–162. doi:10.1109/ISMAR.2008.4637346.
- [68] Jack, R.E., Garrod, O.G., Yu, H., Caldara, R., Schyns, P.G., 2012. Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences* 109, 7241–7244.
- [69] Jia, Y., Zhang, Y., Weiss, R.J., Wang, Q., Shen, J., Ren, F., Chen, Z., Nguyen, P., Pang, R., Moreno, I.L., Wu, Y., 2019. Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis. arXiv:1806.04558 [cs, eess] arXiv:1806.04558.
- [70] Johnson, N.A., Cooper, R.B., Chin, W.W., 2009. Anger and flaming in computer-mediated negotiation among strangers. *Decision Support Systems* 46, 660–672. doi:10.1016/j.dss.2008.10.008.
- [71] Jørgensen, J., 2018. Appeal and Perceived Naturalness of a Soft Robotic Tentacle, in: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM, Chicago IL USA. pp. 139–140. doi:10.145/3173386.3176985.
- [72] Jütten, T., 2011. The Colonization Thesis: Habermas on Reification. *International Journal of Philosophical Studies* 19, 701–727. doi:10.1080/09672559.2011.629672.
- [73] Kabiljo, L., 2019. Virtual Reality Fostering Empathy: Meet the Enemy. *Studies in Art Education* 60, 317–320.
- [74] Kac, E., 2005. *Telepresence and Bio Art : Networking Humans, Rabbits and Robots.* University of Michigan Press, Ann Arbor.
- [75] Kaye, L.K., Malone, S.A., Wall, H.J., 2017. Emojis: Insights, Affordances, and Possibilities for Psychological Science. *Trends in Cognitive Sciences* 21, 66–68. doi:10.1016/j.tics.2016.10.007.
- [76] Keating, E., 2017. Challenges of Conducting Interaction with Technologically Mediated Bodies, in: *Intercorporeality: Emerging Socialities in Interaction.* Oxford University Press, p. 303.
- [77] Kendon, A., 1990. *Conducting Interaction: Patterns of Behavior in Focused Encounters.* volume 7. CUP Archive.
- [78] Kerruish, E., . Affective Touch in Social Robots , 20.
- [79] Kidd, C., Taggart, W., Turkle, S., 2006. A sociable robot to encourage social interaction among the elderly, in: *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006., IEEE, Orlando, FL, USA.* pp. 3972–3976. doi:10.1109/ROBOT.2006.1642311.
- [80] Kouli, E.R., Kontiainen, L., . How Will AI Shape the Future of Law? , 224.
- [81] Kozima, H., Nakagawa, C., 2006. Social robots for children: Practice in communication-care, in: *9th IEEE International Workshop on Advanced Motion Control, 2006., IEEE, Istanbul, Turkey.* pp. 768–773. doi:10.1109/AMC.2006.1631756.
- [82] Kress, G., 2009. *Multimodality: A Social Semiotic Approach to Contemporary Communication.* Routledge.
- [83] Kristoffersson, A., Coradeschi, S., Loutfi, A., 2013. A Review of Mobile Robotic Telepresence. *Advances in Human-Computer Interaction* 2013, 1–17. doi:10.1155/2013/902316.
- [84] Kuwamura, K., Minato, T., Nishio, S., Ishiguro, H., 2012. Personality distortion in communication through teleoperated robots, in: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, IEEE, Paris, France.* pp. 49–54. doi:10.1109/ROMAN.2012.6343730.
- [85] Laban, R., Ullmann, L., 1971. *The Mastery of Movement.*
- [86] Lala, D., Inoue, K., Kawahara, T., 2019. Smooth Turn-taking by a Robot Using an Online Continuous Model to Generate Turn-taking Cues, in: *2019 International Conference on Multimodal Interaction, Association for Computing Machinery, Suzhou, China.* pp. 226–234. doi:10.1145/3340555.3353727.
- [87] Lee, H., Kim, Y.H., Lee, K.k., Yoon, D.K., You, B.J., 2015. Designing the appearance of a telepresence robot, M4K: A case study, in: *International Workshop on Cultural Robotics,* Springer. pp. 33–43.
- [88] Lee, K.M., Nass, C., 2005. Social-Psychological Origins of Feelings of Presence: Creating Social Presence With Machine-Generated Voices. *Media Psychology* 7, 31–45. doi:10.1207/S1532785XMEP0701_2.
- [89] Lehman-Wilzig, S., 2000. The Tower of Babel vs the power of babble: Future political, economic and cultural consequences of synchronous, automated translation systems (SATS). *New Media & Society* 2, 467–494.
- [90] Lombard, M., Ditton, T., 1997. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication* 3. doi:10.1111/j.1083-6101.1997.tb00072.x.
- [91] Lorenz, T., Weiss, A., Hirche, S., 2016. Synchrony and Reciprocity: Key Mechanisms for Social Companion Robots in

- Therapy and Care. International Journal of Social Robotics 8, 125–143. doi:10.1007/s12369-015-0325-8.
- [92] Malafouris, L., 2013. How Things Shape the Mind. MIT Press.
- [93] Maoz, I., 2005. Evaluating the Communication between Groups in Dispute: Equality in Contact Interventions between Jews and Arabs in Israel. Negotiation Journal 21, 131–146. doi:10.1111/j.1571-9979.2005.00050.x.
- [94] Maoz, I., 2011. Does contact work in protracted asymmetrical conflict? Appraising 20 years of reconciliation-aimed encounters between Israeli Jews and Palestinians. Journal of Peace Research 48, 115–125. doi:10.1177/0022343310389506.
- [95] Masuda, M., Kato, S., 2010. Motion rendering system for emotion expression of human form robots based on Laban movement analysis, in: 19th International Symposium in Robot and Human Interactive Communication, IEEE, Viareggio, Italy. pp. 324–329. doi:10.1109/ROMAN.2010.5598692.
- [96] Mehrabian, A., Russell, J.A., 1974. An Approach to Environmental Psychology. the MIT Press.
- [97] Merleau-Ponty, M., 2013. Phenomenology of Perception. Routledge.
- [98] Merleau-Ponty, M., Lefort, C., Lingis, A., 1968. The Visible and the Invisible : Followed by Working Notes. Northwestern University Press, Evanston.
- [99] Messinger, P.R., Ge, X., Smirnov, K., Stroulia, E., Lyons, K., 2019. Reflections of the extended self: Visual self-representation in avatar-mediated environments. Journal of Business Research 100, 531–546. doi:10.1016/j.jbusres.2018.12.020.
- [100] Michaud, F., . Telepresence Robot for Home Care Assistance , 6.
- [101] Minsky, M., 1980. Telepresence .
- [102] Mitchell, W.J., Szerszen, K.A., Lu, A.S., Schermerhorn, P.W., Scheutz, M., MacDorman, K.F., 2011. A Mismatch in the Human Realism of Face and Voice Produces an Uncanny Valley. i-Perception 2, 10–12. doi:10.1068/i0415.
- [103] Monzó-Nebot, E., 2019. Translators and interpreters as agents of diversity. Managing myths and pursuing justice in postmonolingual societies. Translating and Interpreting Justice in a Postmonolingual Age , 9.
- [104] Mori, M., MacDorman, K.F., Kageki, N., 2012. The Uncanny Valley [From the Field]. IEEE Robotics Automation Magazine 19, 98–100. doi:10.1109/MRA.2012.2192811.
- [105] Muri, S.A., 2007. Beyond the face: Art therapy and self-portraiture. The Arts in Psychotherapy 34, 331–339. doi:10.1016/j.aip.2007.05.002.
- [106] Nagendran, A., Steed, A., Kelly, B., Pan, Y., 2015. Symmetric telepresence using robotic humanoid surrogates: Robotic symmetric telepresence. Computer Animation and Virtual Worlds 26, 271–280. doi:10.1002/cav.1638.
- [107] Nakata, T., Sato, T., Mori, T., Mizoguchi, H., 1998. Expression of emotion and intention by robot body movement, in: Proceedings of the 5th International Conference on Autonomous Systems.
- [108] Ng, R., Lindgren, R., 2013. Examining the effects of avatar customization and narrative on engagement and learning in video games, in: Proceedings of CGAMES'2013 USA, pp. 87–90. doi:10.1109/CGames.2013.6632611.
- [109] Nie, J., Pak, M., Marin, A.L., Sundar, S.S., 2012. Can you hold my hand?: Physical warmth in human-robot interaction, in: Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '12, ACM Press, Boston, Massachusetts, USA. p. 201. doi:10.1145/2157689.2157755.
- [110] Noguchi, Y., Tanaka, F., 2020. OMOY: A Handheld Robotic Gadget that Shifts its Weight to Express Emotions and Intentions, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, ACM, Honolulu HI USA. pp. 1–13. doi:10.1145/3313831.3376775.
- [111] Nowak, K., . Defining and Differentiating Copresence, Social Presence and Presence as Transportation , 24.
- [112] Oda, R.K., . Tools and Techniques for Rhythmic Synchronization in Networked Musical Performance , 162.
- [113] Ogawa, K., Nishio, S., Koda, K., Taura, K., Minato, T., Ishii, C.T., Ishiguro, H., 2011. Telenoid: Tele-presence android for communication, in: ACM SIGGRAPH 2011 Emerging Technologies on - SIGGRAPH '11, ACM Press, Vancouver, British Columbia, Canada. pp. 1–1. doi:10.1145/2048259.2048274.
- [114] Pagotto, L., Voci, A., Maculan, V., 2010. The effectiveness of intergroup contact at work: Mediators and moderators of hospital workers' prejudice towards immigrants. Journal of Community & Applied Social Psychology 20, 317–330. doi:10.1002/casp.1038.
- [115] Paynter, H.M., 1961. Analysis and Design of Engineering Systems. MIT press.
- [116] Pelczer, I., Cabiedes, F., Gamboa, F., 2007. Expressions of Emotions in Virtual Agents: Empirical Evaluation, in: 2007 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, IEEE, Ostuni, Italy. pp. 31–35. doi:10.1109/VECIMS.2007.4373923.
- [117] Peled, A., 2019. Soft Robotic Incarnation. Ph.D. thesis. Aalto University.
- [118] Perry, C., 2013. Machine learning and conflict prediction: A use case. Stability: International Journal of Security and Development 2, 56.
- [119] Pettigrew, T.F., 1998. Intergroup contact theory. Annual review of psychology 49, 65–85.
- [120] Pettigrew, T.F., Tropp, L.R., 2006. A meta-analytic test of intergroup contact theory. Journal of Personality and Social Psychology 90, 751–783. doi:10.1037/0022-3514.90.5.751.
- [121] Pettigrew, T.F., Tropp, L.R., Wagner, U., Christ, O., 2011. Recent advances in intergroup contact theory. International Journal of Intercultural Relations 35, 271–280. doi:10.1016/j.ijintrel.2011.03.001.
- [122] Phillips-Silver, J., Aktipis, C.A., A. Bryant, G., 2010. The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement. Music Perception 28, 3–14. doi:10.1525/mp.2010.28.1.3.
- [123] Reich-Stieber, N., Eysel, F., Hohnemann, C., 2019. Involve the user! Changing attitudes toward robots by user participation in a robot prototyping process. Computers in Human Behavior 91, 290–296. doi:10.1016/j.chb.2018.09.041.
- [124] Richardson, D.C., Dale, R., Tomlinson, J.M., 2009. Conversation, Gaze Coordination, and Beliefs About Visual Context. Cognitive Science 33, 1468–1482. doi:10.1111/j.1551-6709.2009.01057.x.
- [125] Saguy, T., Tausch, N., Dovidio, J.F., Pratto, F., 2009. The Irony of Harmony: Intergroup Contact Can Produce False Expectations for Equality. Psychological Science 20, 114–121. doi:10.1111/j.1467-9280.2008.02261.x.
- [126] Schiffrin, D., 1974. Handwork as ceremony: The case of the handshake. Semiotica 12, 189–202.
- [127] Schumann, S., Klein, O., Douglas, K., Hewstone, M., 2017. When is computer-mediated intergroup contact most promising? Examining the effect of out-group members' anonymity on prejudice. Computers in Human Behavior 77, 198–210. doi:10.1016/j.chb.2017.08.006.
- [128] Shafir, T., Tsachor, R.P., Welch, K.B., 2016. Emotion Regulation through Movement: Unique Sets of Movement Characteristics are Associated with and Enhance Basic Emotions. Frontiers in Psychology 6. doi:10.3389/fpsyg.

- 2015.02030.
- [129] Shen, S., Slovak, P., Jung, M.F., 2018. "Stop. I See a Conflict Happening.": A Robot Mediator for Young Children's Interpersonal Conflict Resolution, in: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction, ACM, Chicago IL USA. pp. 69–77. doi:10.1145/3171221.3171248.
- [130] Shnabel, N., Nadler, A., 2008. A needs-based model of reconciliation: Satisfying the differential emotional needs of victim and perpetrator as a key to promoting reconciliation. *Journal of Personality and Social Psychology* 94, 116–132. doi:10.1037/0022-3514.94.1.116.
- [131] Sholt, M., Gavron, T., 2006. Therapeutic Qualities of Clay-work in Art Therapy and Psychotherapy: A Review. *Art Therapy* 23, 66–72. doi:10.1080/07421656.2006.10129647.
- [132] Sievers, B., Polansky, L., Casey, M., Wheatley, T., 2013. Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences* 110, 70–75. doi:10.1073/pnas.1209023110.
- [133] Silva, S., 2013. Reification and Fetishism: Processes of Transformation. *Theory, Culture & Society* 30, 79–98. doi:10.1177/0263276412452892.
- [134] Sorouspour, S., Setoodeh, P., 2005. Multi-operator/multi-robot teleoperation: An adaptive nonlinear control approach, in: 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE. pp. 1576–1581.
- [135] Söeffner, J., Nam, C.S., 2007. Co-presence in shared virtual environments: Avatars beyond the opposition of presence and representation, in: International Conference on Human-Computer Interaction, Springer. pp. 949–958.
- [136] Spears, R., Postmes, T., Lea, M., Wolbert, A., 2002. When are net effects gross products? *Communication. Journal of Social Issues* 58, 91–107.
- [137] Stanovsky, G., Smith, N.A., Zettlemoyer, L., 2019. Evaluating gender bias in machine translation. arXiv preprint arXiv:1906.00591 .
- [138] Steuer, J., 1992. Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication* 42, 73–93. doi:10.1111/j.1460-2466.1992.tb00812.x.
- [139] Stiehl, W., Lieberman, J., Breazeal, C., Basel, L., Lalla, L., Wolf, M., 2005. Design of a therapeutic robotic companion for relational, affective touch, in: ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005., IEEE, Nashville, TN, USA. pp. 408–415. doi:10.1109/ROMAN.2005.1513813.
- [140] Stupacher, J., Wood, G., Witte, M., 2017. Synchrony and sympathy: Social entrainment with music compared to a metronome. *Psychomusicology: Music, Mind, and Brain* 27, 158–166. doi:10.1037/pmu0000181.
- [141] Sycara, K.P., 1993. Machine learning for intelligent support of conflict resolution. *Decision Support Systems* 10, 121–136. doi:10.1016/0167-9236(93)90034-Z.
- [142] Tanaka, F., Takahashi, T., Matsuzoe, S., Tazawa, N., Morita, M., 2014. Telepresence robot helps children in communicating with teachers who speak a different language, in: Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction - HRI '14, ACM Press, Bielefeld, Germany. pp. 399–406. doi:10.1145/2559636.2559654.
- [143] Todo, T., 2018. SEER: Simulative emotional expression robot, in: ACM SIGGRAPH 2018 Emerging Technologies, ACM, Vancouver British Columbia Canada. pp. 1–2. doi:10.1145/3214907.3214921.
- [144] Vázquez, M., Carter, E.J., McDorman, B., Forlizzi, J., Steinfield, A., Hudson, S.E., 2017. Towards Robot Autonomy in Group Conversations: Understanding the Effects of Body Orientation and Gaze, in: 2017 12th ACM/IEEE Interna-
- tional Conference on Human-Robot Interaction (HRI, pp. 42–52.
- [145] Voci, A., Hewstone, M., 2003. Intergroup Contact and Prejudice Toward Immigrants in Italy: The Mediational Role of Anxiety and the Moderational Role of Group Salience. *Group Processes & Intergroup Relations* 6, 37–54. doi:10.1177/1368430203006001011.
- [146] Walther, J.B., 1996. Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication research* 23, 3–43.
- [147] Walther, J.B., Hoter, E., Ganayem, A., Shonfeld, M., 2015. Computer-mediated communication and the reduction of prejudice: A controlled longitudinal field experiment among Jews and Arabs in Israel. *Computers in Human Behavior* 52, 550–558. doi:10.1016/j.chb.2014.08.004.
- [148] Wessel, T., 2009. Does Diversity in Urban Space Enhance Intergroup Contact and Tolerance? *Geografiska Annaler: Series B, Human Geography* 91, 5–17. doi:10.1111/j.1468-0467.2009.00303.x.
- [149] White, F.A., Harvey, L.J., Abu-Rayya, H.M., 2015. Improving Intergroup Relations in the Internet Age: A Critical Review. *Review of General Psychology* 19, 129–139. doi:10.1037/gpr0000036.
- [150] Willemse, C.J.A.M., Toet, A., van Erp, J.B.F., 2017. Affective and Behavioral Responses to Robot-Initiated Social Touch: Toward Understanding the Opportunities and Limitations of Physical Contact in Human–Robot Interaction. *Frontiers in ICT* 4, 12. doi:10.3389/fict.2017.00012.
- [151] Zancanaro, M., Stock, O., Eisikovits, Z., Koren, C., Weiss, P.L., 2012. Co-narrating a conflict: An interactive tabletop to facilitate attitudinal shifts. *ACM Transactions on Computer-Human Interaction* 19, 24:1–24:30. doi:10.1145/2362364.2362372.
- [152] Zhang, S., Zhang, S., Huang, T., Gao, W., 2016. Multi-modal Deep Convolutional Neural Network for Audio-Visual Emotion Recognition, in: Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval - ICMR '16, ACM Press, New York, New York, USA. pp. 281–284. doi:10.1145/2911996.2912051.
- [153] Zheng, C.Y., . Soft grippers not only grasp fruits: From affective to psychotropic HRI .