Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE) Scale

STEEVEN VILLA, LMU Munich, Germany
JASMIN NIESS, University of Oslo, Norway
ALBRECHT SCHMIDT, LMU Munich, Germany
ROBIN WELSCH, Aalto University, Finland

Human augmentation technologies (ATs) are a subset of ubiquitous on-body devices designed to improve cognitive, sensory, and motor capacities. Although there is a large corpus of knowledge concerning ATs, less is known about societal attitudes towards them and how they shift over time. To that end, we developed The Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE) Scale, which measures how users of ATs are perceived. To develop the scale, we first created a list of possible scale items based on past work on how people respond to new technologies. The items were then reviewed by experts. Next, we performed exploratory factor analysis to reduce the scale to its final length of thirteen items. Subsequently, we confirmed test-retest validity of our instrument, as well as its construct validity. The SHAPE scale enables researchers and practitioners to understand elements contributing to attitudes toward augmentation technology users. The SHAPE scale assists designers of ATs in designing artifacts that will be more universally accepted.

CCS Concepts: • **Human-centered computing** → **Ubiquitous computing**; **Ubiquitous and mobile computing design and evaluation methods**; *HCI design and evaluation methods.*

Additional Key Words and Phrases: Human augmentation, attitudes scale, Human-Computer integration, societal perception

1 INTRODUCTION

Human augmentation technologies are technologies that aim to improve human performance to a level that would not have been possible otherwise [3, 39]. These kinds of technologies could change how people interact with their surroundings and do tasks that require specific physical, mental, or sensory skills [32, 58, 68, 73]. Recent advances in artificial intelligence (AI) [4, 85], augmented reality (AR) [1, 69, 78], prosthetics [59, 74], robotics [51, 71], and wearables [61, 65], among other technologies [86], have made the vision of enhancing human skills tangible. However, a significant portion of this research has been technical in nature, has not addressed the social implications of these technologies in depth [58, 68], and was often relying on qualitative methods or not validated questionnaires for evaluation purposes. To allow for systematic research on the social implications of human augmentation technology, a tool is needed that enables the measurement of attitudes toward augmented humans (AHs). Consequently, our work aims to introduce a validated scale for assessing attitudes towards AHs, effectively

---

Authors' addresses: Steeven Villa, steeven.villa@ifi.lmu.de, LMU Munich, Munich, Germany; Jasmin Niess, jasmin.niess@unisg.ch, University of Oslo, Oslo, Norway; Albrecht Schmidt, albrecht.schmidt@lmu.de, LMU Munich, Munich, Germany; Robin Welsch, , Aalto University, Espoo, Finland.

addressing this research gap. The scale provides researchers with a standardized and reliable measurement tool, facilitating the collection of robust quantitative data and enabling comparative analysis across studies and samples.

Attitudes toward human augmentation play a crucial role in the adoption and socially acceptable development of performance-enhancing technologies [82]; the lack of social acceptability of augmentation devices could affect the self-perception of Augmentation Technologies (ATs) users and hinder the adoption of novel technologies [44]. In addition, it can result in stigmatization and unfavorable evaluations, negatively impacting the full realization of the benefits and potential of ATs, for example [56] discovered that individuals referred to as "cyborg" are perceived as less warm. In a broader sense, acceptability and perception of technologies have been extensively discussed in HCI (for a review, see Koelle et al. [42]) and have led to the development of a variety of instruments to measure dimensions such as Technological Readiness [47], Attitudes Towards Artificial Intelligence [72], Creepiness of Technology [84], or adaptations of the Technology Acceptance Model (TAM) [53, 60]. Yet, traditional measurement instruments focus on interfaces or users and fall short in evaluating the perception of integrated systems, such as human augmentation. Consequently, there is a critical need to establish a systematic approach to gauge attitudes toward augmented humans.

In an effort to advance research on attitudes toward augmented humans, we introduce the Society's Attitudes Towards Human Augmentation and Performance Enhancement Technologies (SHAPE). Following a standardized procedure [10], we conducted a concept selection for concepts related to augmented humans, followed by a search for instruments addressing these concepts. After selecting and adapting the items from those instruments to the scope of human augmentation, we empirically evaluated the consistency and coverage of the items in six expert interviews with participants with a proven research track record in the field of human augmentation. We used exploratory factor analysis to reduce the number of items, and confirmatory factor analysis to obtain the final scale, which was then psychometrically validated.

This manuscript reports the development of SHAPE scale. Comprising 13 items, the final scale measures social attitudes toward augmented humans. In the development, we identify two factors that determine attitudes towards augmented humans; Social Threat, attitudes about augmented humans causing harm to individuals and society, and Agency, an attitude describing whether augmented humans have agency over their own actions. We show that the SHAPE scale is reliable by showing it is stable over a time of about 2.5 weeks and that it converges with technology centric-measures, technology readiness index, and stereotype-centric measures, stereotype content model (SCM) and can predict whether people want to adopt human augmentation technology themselves. We, thus, provide researchers and practitioners in the domain of human augmentation with a psychometrically validated and brief measure that can quantify attitudes toward augmented humans.

2 RELATED WORK

To set the stage for our inquiry, we first analyze why considering attitudes towards new technologies and social acceptance concerning new technologies is crucial for their design. Next, we outline the concept of human augmentation. Finally, we discuss the interplay between technologies that improve human performance and societal attitudes and how it differs from the concept of technology acceptance.

2.1 Society's Influence on Emerging Technologies

Society's perception of a technology has the power to shape its trajectory. Negative perceptions can inhibit widespread acceptance and adoption, and the fear of social stigma can prevent early adopters from embracing new technologies [44]. This can result in a vicious cycle in which the lack of early adopters leads to additional negative perceptions of a specific technology, which in turn discourages potential adopters [42]. This is a critical challenge for emerging technologies, as preconceived biases can cloud the objective evaluation of their advantages

and disadvantages, thereby limiting their potential for positive impact [41]. Emerging technologies such as AI [76], Robotics [36], and Human augmentation [73], are particularly susceptible to this issue partly due to their extensive coverage in science fiction literature, movies, the news, and social media [8] as well as because of their closeness to or their resemblance of the human body [25]. These negative attitudes can be attributed to social validation, perceived aesthetics, intrinsic motivation, and technology-related stigma [80]. Collectively, these factors influence an individual's perception in terms of social acceptability and, ultimately, their willingness to adopt a new technology. To illustrate this, users often refuse to accept assistive technologies [75], even though many such tools have been shown to effectively compensate for users' hearing [49], sight [48], and movement [22] impairments. Since technology acceptability has been recognized as a key concern in HCI [42], various instruments for measuring public opinion on technological innovations have been developed. For instance, measurement scales based on the technology acceptance model [53], The WEAR scale [40], and, more recently, the creepy technology scale [84], amongst others. Note that these scales often focus on the technology itself and thus may not apply to cases of human-computer integration, where the lines between technology and the user blur.

As users are increasingly likely to experience augmented humans in their everyday life, HCI needs to understand more about the attitudes toward augmented humans. To this end, we present the SHAPE scale to aid in the development, dissemination and general adoption of the next generation of digital technologies designed to augment human capabilities.

2.2 Human Augmentation Technologies

The use of augmentation technologies can improve human senses, human thinking, and human action. Augmentation technologies can thus be categorized as either sensory, cognitive, or motor augmentations [68]. Initial sensory enhancements arose from the need to compensate for impairments; however, the adoption of those technologies by non-impaired individuals resulted in improved skills. For example, see Proulx [67] for the case of improved hearing or Danilov et al. [21] for the case of improved vision. Motor augmentations, envisioned as technologies to compensate for limited mobility, evolved in a similar fashion and nowadays can augment the user [46]. Thus, the origins of human augmentation can be traced back to the creation of aids for people who required assistive technologies. [32, 37]. Researchers in the field of human augmentation nowadays are interested in exploring the possibility of augmenting abilities beyond human limitations leveraging the latest developments of digital technologies [38, 73]. Humans equipped with exoskeletons, for example, would be able to lift significantly more weight than they could before [14, 15, 54]. New body conceptions combined with technological advancements can enable humans to move in ways that would not otherwise be possible [59]. Life-logging devices support memory to allow for remembering experiences more vividly and longer [45]. Firefighters can use the combination of augmented reality headsets and thermal cameras to allow them to perceive the infrared spectrum which is helpful when working in high-temperature environments [1, 2]. Thus, human augmentation's objective is not only to re-enable users with impairments but also to extend a user's abilities beyond human limitations.

2.3 Social Attitudes Towards Performance-Enhancing Technologies

Psychology and medicine have extensively studied attitudes towards performance-enhancing non-digital technologies and found that they vary greatly depending on the context [18]. For instance, the use of performance-enhancing drugs in sports has been a source of disagreement for many years. It has been reported that society, at different layers, have markedly different attitudes towards enhancing supplements, depending on their social affiliations to specific groups [12]. Dijkstra and Schuijff [24] found a widespread attitude of mild disapproval to strong disapproval of using enhancement technologies for applications other than medical treatment. Moreover, their findings suggest that the acceptability of enhancement use is dependent on the motivation behind it, with

socially motivated enhancements being perceived more positively than those used for personal gain. The debate around doping in sports has greatly contributed to the search of strategies and tools to measure the attitudes toward performance-enhancing technologies and create an understanding of their impact; reflected in the creation of tools like the Performance Enhancement Attitudes Scale (PEAS) [64]. Yet a big part of the efforts to measure attitudes towards performance enhancement from the medical, psychological and sport science domain have been directed towards predicting doping behavior by connecting the attitudes and the chances that it can be correlated to the doping behavior intention and use [30].

In contrast, attitudes towards digital technology enhancements bring new challenges compared to non-digital enhancements. For example, the use of anabolic steroids is generally seen as a punishable behavior [64], while using exoskeletons to increase strength is seen as something needed in some cases [26]. Or even more, pose difficult situations with no obvious consensus, such as amputee athletes outperforming their non-amputee peers [5]. Consequently, there is a need for a tool to measure attitudes towards human augmentation technologies, as they continue to evolve and shape the trajectory of technological advancements.

3 SCALE FORMATION

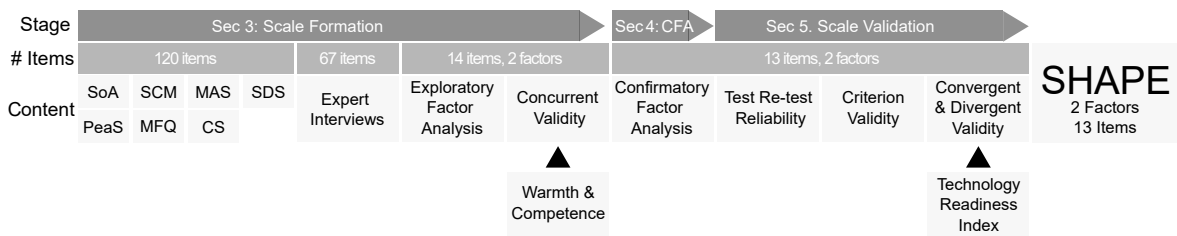Fig. 1. Study diagram:In the first stage, Scale Formulation, we searched the literature for intersecting instruments and generated an initial set of items. We then reduced the number of items using expert interviews, and finally, we performed an exploratory factor analysis to reduce dimensionality and discover the underlying structure of the factors. The construct's structure was then assessed using a confirmatory factor analysis. Finally, we ran a series of tests to validate the SHAPE scale psychometrically to establish and validate its final structure.

The SHAPE scale was developed with the aim of facilitating standardized measurement of attitudes towards human augmentation and performance-enhancing technologies in the field of human-computer interaction (HCI). The SHAPE scale will enable HCI designers to create human augmentation and performance-enhancing technologies that are better aligned with the attitudes and expectations of the general public, thus promoting wider social acceptance and adoption of ATs.

This study has been approved by the Institutional Review Board of the University of [Anonymized] (approval number [Anonymized]).

3.1 Item Generation

While several instruments for evaluating perceptions and attitudes towards various technological instances exist [18, 84], a gap remains in the assessment of attitudes towards technologies that blur the boundaries between humans and machines [58, 82] as for example Electrical Muscle Stimulation (EMS) to improve reaction times [39], or the use of wearable robotics to control multiple supernumerary limbs at the same time [71]. To address this gap, we constructed the SHAPE scale. As a first step, we conducted an analysis of existing studies and measures in related research fields. This analysis aimed to synthesize the data from instruments with intersecting concepts, such as the sense of agency [81], or attitudes towards assistive technology users [27], and inform the development

of the SHAPE scale. The items selected from these instruments were then adapted and grouped to form the initial pool of items for the SHAPE scale. Here, we describe the concepts and instruments selected to design this initial pool of items.

3.1.1 Sense of Agency Scale (SoA). The relationship between body and action ownership is fundamental to the formation of our self-perception and perception of others [58]. With the proliferation of augmentation technologies, there is an increasing concern that these tools may alter our sense of self and others [58]. Prior research has shown that SoA is especially important in the context of augmented humans, as ATs may alter an individual's SoA and the amount of effort users invest in a task [58, 82].

SoA can be defined as the subjective experience of initiating and controlling one's own actions [57]. It is typically measured by means of self-report through the Sense of Agency Scale [81]. The items extracted from the sense of agency (SoA) scale evaluates an individual's perceived control over their body and actions, providing valuable insights into their subjective experience of agency. Given that the items on the SoA scale are framed in the first person, we modified them to reflect a reference to a third-person perspective, e.g., the item "I am in full control of what I do" changed to "An augmented human is in control of what they do."

3.1.2 Social Stereotype - Stereotype Content Model (SCM). The SCM is a psychological theory that explains how individuals develop stereotypes about others. It describes stereotypes regarding distinct social groups along two broad dimensions: Warmth and Competence [28, 29]. These factors allow for the prediction of a range of emotions and perceptions, including pride, pity, contempt, and envy towards a distinct social group (in this case towards augmented humans). In HCI, the SCM has been used to describe labeler bias [33], people stereotypes for artificial intelligence systems [55]. In line with [55], we adapted the items of [29] to measure how attitudes vary according to perceptions of competence and warmth for augmented humans. For example, an adapted item would be phrased as follows "In general, augmented humans are perceived as warm."

3.1.3 Multidimensional Attitudes Scale Toward Persons with Disabilities (MAS). The manner in which an augmented human uses the augmentation technology, the reasons for its use, and whether or not the individual had a prior disability, are among the factors that can evoke emotions in the observer [82].

The MAS Scale is a validated measure of attitudes toward people with disabilities. It provides a comprehensive picture of the attitudes of society towards this population. The MAS has been shown to be a valid and reliable measure of attitudes toward people with disabilities [27]. It can be used to identify areas where attitudes may need to be altered and to inform the design of interventions aimed at reducing prejudice and promoting the inclusion of persons with disabilities. The items from the MAS scale provide valuable information about the observer's emotions, thoughts, and behaviors towards ATs users.

3.1.4 Social Distance Scale (SDS). The SDS [11] is constructed to measure stigma and is a routine measure in stigma research [19]. As Augmented individuals may face stigma and discrimination, we adapted these items to study how social stigma may affect augmented humans. The original SDS Scale measures how far away from a group (such as people with a specific disease) an individual would like to remain. We adapted the SDS items to fit into the human augmentation context, resulting in items with the following structure "I would have an augmented human as citizens in my country"

3.1.5 Performance Enhancement Attitudes Scale (PeaS). PeaS [30]focuses on measuring attitudes toward performance enhancement via non-technological means, such as doping. Given the similarities between human augmentation and performance enhancement, this group of items adapts key PEAS components to the context of human augmentation. Items such as "Doping is necessary to be competitive." were adapted to the human augmentation context, resulting in "An augmented human is more competitive."

3.1.6 Moral Foundations Questionnaire (MFQ). Numerous factors, such as sociocultural context and individual personality traits influence the perceptions of morality. We adapted the MFQ [31] to evaluate how the observer integrates the concept of human augmentation into their personal values, cultural norms, and political ideologies. The MFQ quantifies moral convictions. It assesses an individual's moral sensitivity across five theoretical dimensions, including Harm/Care, Fairness/Reciprocity, Ingroup/Loyalty, Authority/Respect, and Purity/Sanctity. We extracted and adapted the items relevant to human augmentation. e.g., the item "Whether or not someone violated standards of purity and decency" from the MFQ is reflected in the initial pool of items as " An augmented human would violate standards of purity and decency."

3.1.7 Cross-Sectional Studies in Human Augmentations (CS). Recent research in human-augmentation and integration technologies has uncovered a variety of factors that influence society's evaluation of adopters of these technologies [58, 70, 82]. To better understand how people perceive human augmentation technologies and the reasons behind their assessment, Villa et al. [82] conducted a mixed-method cross-sectional study. They discovered that seven factors influenced people's opinions of human-augmentation users: Privacy, Peril, Ownership, Motivation, Perception of Achievements, and Personal Preference. We added and modified these items to the original set of SHAPE scale items. e.g, the item "I think this person has to disclose the presence of this augmentation in their body to other people." was adapted to be depersonalized as follows "An augmented human has to disclose his augmentation."

3.1.8 Item Reduction. To construct a coherent and consistent initial item set based on the instruments described above, the authors have put forth a set of criteria that would inform the wording and selection of the items. In detail, it was prioritized the use of positive, unambiguous and concise phrasing, the use of depersonalized and hypothetical language whenever possible [10], the use of unemotional language, avoidance of abbreviations and that no prior knowledge is needed for the respondent.

One researcher initially reformulated the initial items according to the established criteria. Subsequently, two researchers separately evaluated the wording of the items independently. A final discussion was then held to address and resolve any disagreements regarding the wording of the items. For all items, a seven-point Likert scale was used to measure agreement (7. Very Much) or disagreement (1. Not at all). In this step, we obtained a total of 120 items

3.2 Expert Review

In the subsequent phase, we obtained feedback from six experts who have a record of publication in the domain of human augmentation. The experts provided feedback on each item and suggested eliminating/adding items. Following the expert review, two researchers consolidated and integrated their feedback.

3.2.1 Participants. We invited six experts in human augmentation to participate in the study. Table 1 presents the demographic information of the participants. Experts were selected based on publication-record in the field of human augmentation in the Conference on Human Factors in Computing Systems (CHI) and the Augmented Humans (AHs) conference. A total of seven experts were contacted via email, from whom six accepted the call for participation. The interviews and analysis were performed by two researchers, each researcher interviewed three experts. The interviews took place in a period of approximately one month given the availability of the experts. The experts' participation in the study was strictly voluntary and without financial compensation.

3.2.2 Procedure. Prior to the interview, the experts were provided with a document containing the initial pool of items to become familiar with the content of the scale. During the interview process, the experts were requested to give feedback on the current set of items, propose new items, and modifications or removal. The interviewers

Table 1. Participants' demographic information: Expert review

Expert	Research	Background
E1	Virtual Reality	Computer Science
E2	Information Transfer	Psychology
E3	Physiological Sensing	Computer Science
E4	Thermal Imaging	Human-computer Interaction
E5	Physiological-based Systems	Computer Science
E6	Neuroscience	Cognitive Science

went through each of the 120 items and asked the experts to provide verbal feedback and annotations on the provided document. The annotated documents and interviewer notes were then collected for further analysis.

3.2.3 Analysis. Two researchers participated in the analysis; first, the items suggested for removal by at least one expert were excluded, and the remaining set of items, including those suggested for rephrasing, were discussed. Afterward, the interviewers assessed each item individually and rated the item quality based on the expert feedback on a scale of 1 to 10. Items with high scores (above 6) were retained, items with scores below 3 were excluded, while items with scores between 3 and 6 or were discussed and kept or removed after reaching a consensus.

The expert review started with an initial pool of 120 items sourced from the previously described instruments. The integration of the expert feedback resulted in a reduction of the item pool to 67 partly reformulated items.

3.3 Survey #1

In the next stage of our scale development process, we designed a Qualtrics-based online survey to collect data from participants and conducted an exploratory factor analysis and item reduction. Boateng et al. [10], referring to Comrey [17], recommends a sample size of a minimum of 200 participants for studies of this kind and we exceeded this minimal sample size recommendation with a sample size of $n = 302$ participants.

3.3.1 Participants. The sample was composed of 149 female and 153 male participants with a mean age of 44.4 years (SD=13.0). No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. Participants were recruited through the UK-based platform Prolific, with the sample being drawn from the United Kingdom and the United States. All participants reported English as being their mother tongue. Participation was voluntary and compensated by 9 GBP per hour. The participants were informed that the collected data would be anonymized prior to processing. The survey was distributed in an online format and took participants an average of eight minutes to complete ($M = 8.02$, $SD = 4.24$).

3.3.2 Survey Structure. The survey started with an informed consent form, and after participants gave their consent, they read a scenario depicting the journey of an augmented human interacting with a group of people. The scenario was developed based on Findler et al. [27] and Villa et al. [82] work. This scenario was designed to elicit a range of attitudes towards augmentations by incorporating all possible permutations of cognitive, sensory, and motor augmentations. The following is the scenario:

> Michael went out for lunch with friends to a coffee shop. A man with some technological modifications, with whom Michael is not acquainted, enters the coffee shop and joins the group. Michael is introduced to this person. During the chat, the man tells them that he replaced some of his healthy body parts and replaced them with improved artificial ones: an artificial eye to augment his vision beyond the normal range. Artificial legs to run faster and jump higher than ordinary humans. Additionally, he

got a brain implant to think faster and have more memory than ordinary humans. Shortly after that, everyone else leaves, with only Michael and the man with the technological modifications remaining alone together at the table. Michael has 15 minutes to wait for his ride home.

After this scenario, the participants were presented with a quasi-randomized set of 67 items. Once the participants had responded to all of the questions, their demographic information was collected and the survey concluded.

3.4 Exploratory Factor Analysis

For the item analysis, we inverted the negatively worded items, then we examined the densities of all items and eliminated those with high skew and kurtosis. Then, we conducted a Kaiser-Meyer-Olkin (KMO) factor adequacy test, which evaluates the data's suitability for factor analysis. In a KMO test, values close to 1.0 are desired, and our dataset produced KMO Measure of Sampling Adequacy (MSA) = 0.95. Subsequently, we conducted a Bartlett's Test of Sphericity to evaluate the null hypothesis that the inter-correlations among the variables in the dataset are equal to zero, thereby eliminating the possibility of an identity matrix and ensuring that the variables are suitable for factor analysis ($\tilde{\chi}^2(741) = 9953.585$).

We then performed an exploratory factor analysis. The exploratory factor analysis is a statistical procedure that allows determining the number of underlying factors that explain the pattern of correlation of items [79].

Then, we employed parallel analysis [35] and scree plot analysis [16] to determine the optimal number of underlying factors in the data. The inspection of the scree plot indicated that a two-factor solution was the optimal solution which amounts to extracting factors with an eigenvalue > 1.83.

We then used varimax rotation similar to Woźniak et al. [84]. A varimax rotation produces independent factors; it is an orthogonal rotation method used in factor analysis to maximize the variance of the variable factor loadings while minimizing the number of variables with high factor loadings [87].

From this model, we eliminated all items with loadings below 0.40 and those that were loaded on multiple factors. We merged items with high similarity as a final step. The scale encompassed fourteen items distributed in two factors; seven items per factor. The model had a good fit, KMO MSA = 0.85, Tucker Lewis Index of factoring reliability TLI = 0.842, and, $RMSEA$ = 0.104. Table 2 presents the results of exploratory factor analysis. The first factor is related to Social Threat (ST) [6], indicating that augmented humans will pose a threat to oneself and society. The second factor, Agency (AG), is characterized by a focus on control and includes items that assess the perceived agency of the augmented human over their augmentation. Internal consistency as indicated by Cronbach's alpha was α = 0.852 for ST and α = 0.834 for AG and thus can be regarded as good internal consistency of the scales [20].

3.5 Content Validity

Warmth and competence have been used to structure stereotypical attitudes towards human augmentation [56]. To establish that our novel measure relates to an established measure, we have correlated the ST-scale and the AG-scale to each warmth and competence of the SCM. We observed that the perceived warmth correlates with both SHAPE factors, meaning that a decrease in perceived threat and an increase in control of ATs users increase the perceived warmth. Similarly, we found that competence correlates with both ST-scale and AG-scale control factor, see Table 3. This indicates that an increase in the perceived control over the augmentation and a decrease in the perceived threat increases the perceived competence. These results are consistent with the findings of Meyer and Asbrock [56].

Table 2. The revised version of the SHAPE scale consisted of fourteen items grouped in two factors: Social Threat (ST) and Agency (AG), with item loadings and their respective sources reported.

Item	ID	Source	Factor	
			ST	AG
An augmented human is a threat to society.	I1	CS	0.85	
An augmented human would be dangerous.	I2	CS	0.84	
An augmented human is intimidating.	I3	MAS	0.64	
An augmented human would conform to the traditions of society.	I4	MFQ	0.58	
An augmented human has to disclose their augmentation.	I5	CS	0.58	
An augmented human would do something cruel	I6	MFQ	0.55	
An augmented human is more competitive than a non-augmented human.	I7	PeaS	0.53	
The actions of the augmented human do not match their intentions.	I8	SoA		0.70
An augmented human is not the author of their own actions.	I9	SoA		0.68
An augmented human is just an instrument of something or somebody else.	I10	SoA		0.66
An augmented human does things without any intention.	I11	SoA		0.66
An augmented human suffering through their augmentation should get help.	I12	MFQ		0.50
If an augmented human were to suffer through their augmentation, I would have compassion.	I13	MFQ		0.48
An augmented human is in full control of what they do.	I14	SoA		0.47

Table 3. Correlations between the SHAPE scale factors, Social Threat (ST) and Agency (AG), and the Warmth and Competence scale. degrees of freedom for all the tests are $df = 300$

Factor	ST			AG		
	r	t	p	r	t	p
Warmth	-0.581	-12.392	< .005	-0.491	-9.766	< .005
Competence	-0.205	-3.634	< .005	-0.392	-7.380	< .005

4 SCALE VALIDATION

After building the factor structure of the scale, we continued with the evaluation of the SHAPE scale. We performed a confirmatory factor analysis to test the fit of the structure to novel data. Subsequently, various correlational tests were conducted to assess the scale's content validity and reliability. In this section, we report the first version of the SHAPE scale and evaluate its consistency. We then refine the scale and construct its final version.

4.1 Survey #2

We designed a Qualtrics-based online survey to collect data from participants and conducted a confirmatory factor analysis (CFA) during this phase of the research. It is important to note that the structure of the questionnaire at this stage is identical to that described in subsection 3.3, with the exception that the set of items has been replaced with those obtained from the exploratory factor analysis described in subsection 3.4.

4.1.1 Participants. For this stage, we recruited a sample of $n = 297$ participants, in accordance with the recommendations by Comrey [17] that posit confirmatory factor analysis requires at least 200 participants. The sample consisted of 150 females and 147 males with a mean age of 44.4 ($SD = 13.9$) years. No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. The sample was composed of individuals from the United Kingdom and the United States who were recruited through the British platform Prolific. All participants were native English speakers. The participants were compensated with 9 GBP per hour. All participants were informed of the voluntary nature of their participation and provided with the option to withdraw at any time if they felt uneasy. Participants were also informed that the collected data would be anonymized prior to processing. The survey was distributed online and took respondents an average time of three minutes to complete ($M = 3.45$, $SD = 1.86$).
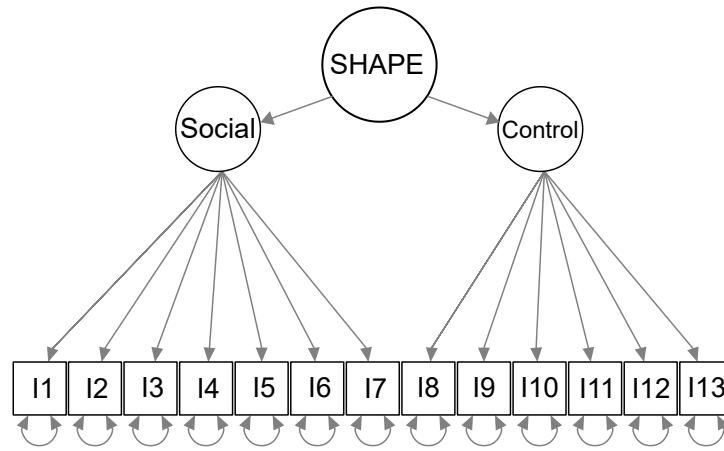
4.2 Confirmatory Factor Analysis (CFA)

Fig. 2. The findings of the confirmatory factor analysis indicated a two-factor model for the SHAPE scale, comprising two inter-correlated subscales.

In order to assess the validity of the SHAPE scale's structure, we conducted a Confirmatory Factor Analysis (CFA). This statistical procedure allowed us to confirm the dimensionality of our proposed factor model. The solution had two factors, see again Table 2. The results of the model fit assessment indicated a sub-optimal fit, as evidenced by the Root Mean Square Error of Approximation (RMSEA) value of greater than 0.1, a Comparative Fit Index (CFI) of 0.93, and a Standardized Root Mean Square Residual (SRMR) of 0.08. Detailed examination of the data revealed high correlations between two items in the Agency factor (items I12 and I13). Also, item I13 was identified as dissimilar due to its wording and was removed to improve the coherence of the SHAPE construct.

We conducted another CFA using the reduced set of items. The confirmatory factor analysis revealed an RMSEA of 0.08, which falls within the acceptable range. Additionally, the CFI was calculated to be 0.97 and the SRMR was determined to be 0.063. Both values are within the desirable bounds, with CFI values above 0.95 and SRMR values below 0.08 being considered indicative of a good fit of the data to the model. Again Cronbach's Alpha for ST was $\alpha = 0.808$ and for AG was $\alpha = 0.809$ can be deemed a good level of internal consistency.

The final compositions of the factors is reported in Table 4 and visualized in Figure 2; A two-factor model consisting of thirteen item; 6 for ST, 7 for AG.

Table 4. The final version of the SHAPE scale consisting of thirteen items. Internal consistency Cronbach's alpha values are displayed on top of their respective item group. Each item is answered in a 7 point Likert scale ranging from (1) Not at all to (7) Very Much. (*) denotes that the item is inverted.

Item	ID	Source
Social Threat $\alpha = 0.808$		
An augmented human is a threat to society.	**S1**	CS
An augmented human would be dangerous.	**S2**	CS
An augmented human is intimidating.	**S3**	MAS
(*) An augmented human would conform to the traditions of society.	**S4**	MFQ
An augmented human has to disclose their augmentation.	**S5**	CS
An augmented human would do something cruel	**S6**	MFQ
(*) An augmented human is more competitive than a non-augmented human.	**S7**	PeaS
Agency $\alpha = 0.809$		
The actions of the augmented human do not match their intentions.	**S8**	SoA
An augmented human is not the author of their own actions.	**S9**	SoA
An augmented human is just an instrument of something or somebody else.	**S10**	SoA
An augmented human does things without any intention.	**S11**	SoA
An augmented human suffering through their augmentation should get help.	**S12**	MFQ
(*)An augmented human is in full control of what they do.	**S13**	SoA

5 TEST-RETEST RELIABILITY AND CONSTRUCT VALIDITY

In this step, we evaluated the construct validity of the SHAPE scale through three methods: (1) Reliability: conducting a test-retest reliability study. (2) Content validity: analyzing the correlation between the SHAPE scale and the willingness to acquire an augmentation, and (3) Convergent validity and discriminant validity: examining the correlation between the SHAPE factors and subscales of the Technology Readiness Index (TRI) that bears subscales that conceptually relate to our measure and subscales that do not [72].

5.1 Data Collection

To gather data and evaluate the three aforementioned points, we developed two online surveys using Qualtrics software. The surveys included the final thirteen items of the SHAPE scale and were completed by a total of n = 103 participants in the first round and n = 78 participants in the second round. The surveys were distributed with a minimum interval of 15 days between assessments ($M = 16.52$, $SD = 0.63$, $min = 15.66$, $max = 18.31$).

5.1.1 Survey #3: Test-Retest first sample, Technology Readiness Index (TRI), and Willingness to acquire an augmentation. The survey started with an informed consent process; following this, participants viewed the same scenario from the first survey (see subsection 3.3 for details). Participants were then presented with the thirteen-item SHAPE scale, and upon completion, participants were asked a binary question regarding their willingness to acquire an augmentation, "I would like to get an augmentation for myself," with response options of "Yes" or

"No." Finally, we administered the Technology Readiness Index (TRI) before concluding the survey by collecting demographic data.

5.1.2 Participants. For this stage, we recruited a sample of n = 103 participants using Prolific, The sample consisted of 51 females and 52 males with a mean age of 45.5 (SD = 13.1) years. No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. The recruiting, compensation and consent scheme were similar to the previous two studies. The survey was distributed online and took respondents almost six minutes to complete ($M = 5.88$, $SD = 3.15$).

5.1.3 Survey #4: Test-retest second sample. Most of the questions from Survey #3 were re-invited to Survey #4, with the Technology Readiness Index (TRI) being the only exception.

5.1.4 Participants. About 80% responded again, n = 78, using Prolific. The sample consisted of 44 females and 34 males with a mean age of 47.3 (SD = 13.9) years. No participants chose not to reveal their identity, and no participant self-identified as non-binary or other. The compensation and consent scheme was the same as in the previous study. The survey was distributed online and took respondents an average of four minutes to complete ($M = 4.19$, $SD = 6.80$).

5.2 Test-retest Reliability

Temporal stability refers to the ability of a scale to produce consistent results when administered to the same participants at different time points [10]. We conducted a test-retest reliability evaluation to assess the temporal stability of the SHAPE scale construct. This psychometric evaluation is commonly used in the scale development process (e.g. Bentvelzen et al. [7], Woźniak et al. [84]) to estimate reliability based on temporal stability.

Similar to Woźniak et al. [84], we calculated calculated a two-way Single-measurement intraclass correlation coefficient (ICC) for consistency and agreement. The ICC quantifies the degree of agreement between two or more continuous measures, values close to 1 indicate a perfect agreement whilst values close to 0 indicates no agreement at all. The ICC, for each subscale[1], indicated good reliability for ST and AG in terms of consistency (ST $\kappa = 0.735$, AG $\kappa = 0.715$) and agreement (ST $\kappa = 0.736$, AG $\kappa = 0.709$), see also Table 5. Additionally we computed Spearman correlations for each subscale, indicating a high correlation between samples; namely for AG we found that $r_s = 0.68$, $p < .005$, and for ST $r_s = 0.707$, $p < .005$.

Table 5. The two-way single-measurement intraclass correlation coefficients (ICC) calculated for both the Social Threat (ST) and Agency (AG) factors of the SHAPE scale.

Factor	ST			AG		
	κ	F	p	κ	F	p
Consistency	0.735	$F(76,76) = 6.54$	$< .005$	0.715	$F(76,76) = 6.02$	$< .005$
Agreement	0.736	$F(76,76.8) = 6.54$	$< .005$	0.709	$F(76,74.2) = 6.02$	$< .005$

To further determine the absolute reliability of the SHAPE scale, we analyzed the data using the Bland and Altman method [9]. Each participant's mean difference between the initial test and the retest was plotted as a function of the means of both test sessions using Bland-Altman plots. The dashed horizontal lines in the plots represent the limits of agreement, which correspond to the 95% confidence interval surrounding the mean

---

[1]We calculated the Intraclass Correlation Coefficient (ICC) for each subscale by averaging the responses of the Likert items on the subscale. Therefore it is important to note that this computation was performed on interval data, as opposed to ordinal data.

difference between the test sessions. These limits indicate the range within which 95% of the values are likely to fall [9, 83].
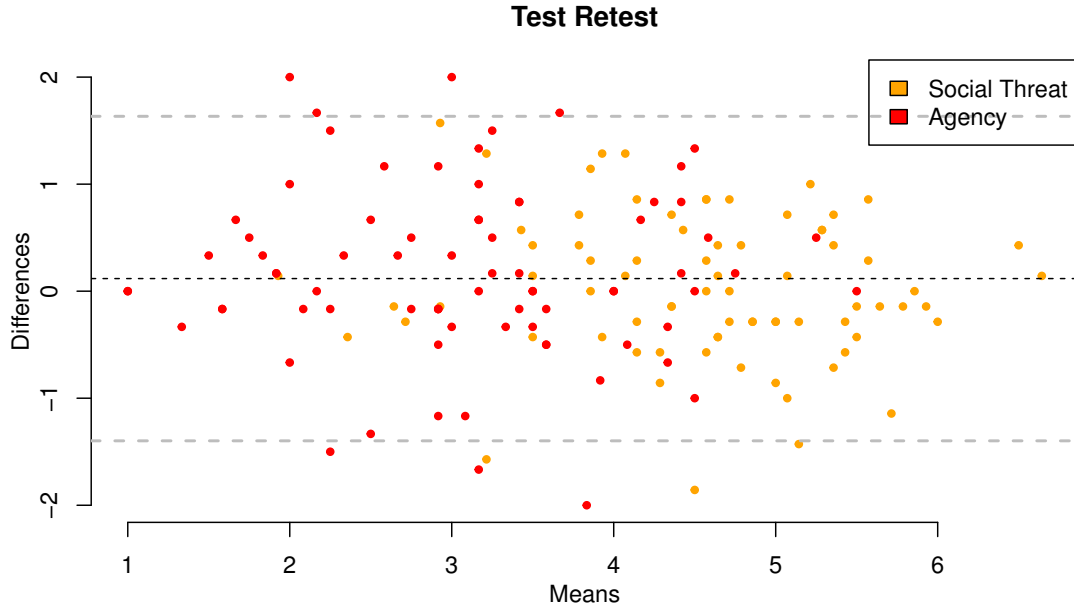

Fig. 3. Bland and Altman plots: difference in SHAPE scores obtained from two surveys (Described at the beginning of this section) as a function of the average score of both test sessions for individual participants, the data is segregated based on ST and AG categories. The mean bias is indicated by the black line, while the 95% limits of agreement are represented by the gray lines.

In the plot (see Figure 3), the mean difference close to zero, (dotted line) indicates that the SHAPE scale has absolute temporal stability on average and the distribution around zero is indicative of reliability not being related to the mean score, thus, demonstrating that it can be reliably administered at different time points and is suitable for use in between-groups or repeated-measures designs.

5.3 Concurrent Validity

In this step, we wanted to investigate the extent to which the factors of the SHAPE scale could predict an individual's inclination to obtain ATs to show concurrent validity. We measured this inclination in Survey one with the response options of "yes" or "no" to the question "I would like to get an augmentation for myself.". We calculated Spearman correlation, for the ST and AG and the above-mentioned question. We found a negative association for the ST, r_s = -.40, p < .001, and AG, r_s = -.31, p = .001, concerning their indication of willingness to acquire an AT. The less threat and the more control they attribute to augmented humans in general, the more likely participants are to indicate they would want to use an AT themselves.

5.4 Convergent & Divergent Validity

Utilizing a methodology similar to that of Schepman and Rodway [72], we assessed the convergent validity of the SHAPE scale by applying the Technology Readiness Index (TRI). The TRI scale comprises 18 items and is frequently used due to its sound psychometric properties [47]. The TRI scale has four subscales: Innovativeness,

Optimism, Discomfort, and Insecurity. The scale has demonstrated the ability to predict user interactions with technology products [72]. The Innovativeness sub-scale is correlated with the tendency to be a thought leader, Optimism with a positive view about technology, discomfort, with the feeling of being overwhelmed by technology and Insecurity, with distrust in technology. We expect Innovativeness to be conceptually independent of ST and AG and discomfort and insecurity to overlap with ST and AG .

To evaluate the internal consistency of the TRI, we determined the Cronbach alpha for each sub-scale. The resulting alpha coefficients were $\alpha = 0.813$ for Innovativeness, $\alpha = 0.698$ for Optimism, $\alpha = 0.725$ for Discomfort, and, $\alpha = 0.792$ for Insecurity. The obtained metrics reflect an acceptable to good performance for each sub-scale. We then obtained the sub-scale values by computing the average of the corresponding items.

The correlations of the SHAPE scale and the TRI factors are presented in Table 6. The correlation analysis indicated that the Social Threat and Agency factors of the SHAPE scale were strongly correlated with the Discomfort and Insecurity scales of the TRI. The less Discomfort and Insecurity experienced in response to technological advancement, the less they perceived augmented humans as threatening and the more control they attributed to them. Thus, we can show convergent validity to negative aspects of technology readiness concepts.

Table 6. Correlation with Technology Readiness Index. degrees of freedom for all the test are $df = 101$

Factor	ST			AG		
	r	t	p	r	t	p
Innovativeness	-0.119	-1.206	0.230	-0.132	-1.340	0.183
Optimism	-0.153	-1.559	0.122	-0.272	- 2.848	< 0.01
Discomfort	0.277	2.902	< 0.005	0.410	4.523	< 0.005
Insecurity	0.213	2.197	< 0.05	0.468	5.326	< 0.005

6 VALIDATION OF SHAPE SCALE IN THE CONTEXT OF DISABILITIES

As a final step, we further explored the fit of the two-factor thirteen item structure of the SHAPE scale to the assessment of augmentation technologies when the user of such technology is an individual with a previous disability condition. We conducted a new Confirmatory Factor Analysis with a modified vignette to reflect a scenario where the technology user is enhancing their skills to compensate for a disability.

6.1 Survey #5

We developed a new online survey using Qualtrics software. The survey included the final structure of the SHAPE scale and was completed by a total of $n = 216$ participants, in accordance to Comrey [17]. The sample consisted of 123 females 91 males and 2 individuals that preferred not to disclose their gender, the mean age of participants was 43.9 ($SD = 11.66$) years. No participants self-identified as non-binary or other. The sample was comprised of individuals from the United Kingdom and the United States who were recruited through the platform Prolific. All participants were native English speakers. Participants were compensated with 9GBP per hour. All participants were informed of the voluntary nature of their participation and provided with the option to withdraw at any time without the need for further explanation. Participants were informed about the data collection and the anonymization policy prior to processing. The survey was distributed online and took respondents an average time of three minutes to complete ($M = 3.55, SD = 1.82$).

6.2 Confirmatory Factor Analysis for the Disabilities Context

We performed a new Confirmatory Factor Analysis utilizing a modified vignette that accounts for a scenario where users are augmenting their skills to offset disability-related impairments instead of augmenting to increase their skills. The original two-factor thirteen items model revealed a sub-optimal fit to this new scenario; the Root Mean Square Error of Approximation (RMSEA) value was greater than 0.1 and the Comparative Fit Index (CFI) had a value of 0.96 with a Standardized Root Mean Square Residual (SRMR) of 0.06. Given these values, we examined the correlations between items using modification indices [50] , which indicate a potential change in χ^2 when adding or removing items, We subsequently deleted S2, S11, and, S12 which were highly skewed and thus had a reduced variance for a disabilities context (more positively valenced responding patterns).

We then performed a CFA with these three items removed. The new analysis revealed an RMSEA of 0.07 which can be considered as a reasonably good fit. For the CFI, we found a value of 0.98, with a SRMR value of 0.043. Both values being considered as a good fit of the data to the model. The Cronbach's Alpha for ST $\alpha = 0.773$ and for AG $\alpha = 0.802$ which can be interpreted as a good level of internal consistency.

In addition, we compared the reduced scores of Survey 5 (validation for disability scenarios) and Survey 3 (test-retest on the initial sample) for both subscales using an unpaired t-test. The results indicated that the score for the non-disability scenario was significantly higher overall (M_{nd} = 3.87, SD_{nd} = 0.76) than the disability scenario (M_d = 3.40, SD_d = 0.74, $t(192.21)$ = 5.10, $p < 0.005$). In addition, the subscales exhibited a similar pattern, for example for Social Threat we found (M_{ST-nd} = 4.19, SD_{ST-nd} = 0.91) and (M_{ST-d} = 3.63, SD_{ST-d} = 0.90, $t(196.91)$ = 5.25, $p < 0.005$), while for Agency we found (M_{AG-nd} = 3.54, SD_{AG-nd} = 0.83) and (M_{AG-d} = 3.18, SD_{AG-d} = 0.82, $t(194.31)$ = 3.57, $p < 0.005$).

Based on the findings of this confirmatory factor analysis, the scoring system for the context of disabilities is described more in detail in subsection 7.1.

7 DISCUSSION

In this section, we provide an overview of our approach, the necessary details for administering the SHAPE scale as well as information on how to use it. In addition, we discuss the limitations of our approach and opportunities for further developments.

In this paper, we introduce the development and validation process of a brief 13-item measure, the SHAPE scale, which was designed to measure attitudes towards humans using digital technologies that enhance human abilities.

We identified a two factorial structure that encompassed attitudes that we summarized under the Social Threat factor, which measures threat to oneself and others, as well as a factor that we summarized under the Agency factor, which describes agency and support for augmented humans. The SHAPE sub-scales were validated and refined in confirmatory factor analysis, showing a good fit based on several fit indices, excellent internal consistency and good test-retest reliability. Also, medium test-retest reliability indicates that attitudes toward augmented humans might be susceptible to changes over time and can thus be used to investigate how attitudes toward augmented humans evolve in the future.

We have evaluated the validity of SHAPE across studies. In Survey #1, we could show that threat and competence relate to the Stereotype-content model; people that attribute low threat to augmented humans perceived them as warmer, while competence of augmented humans was increased for low social threat and more control. This aligns with the findings of Meyer and Asbrock [56], who discovered that individuals with bionic prostheses were perceived as competent without a reduction in perceived warmth.

On the other hand, in Survey #3, we demonstrated construct validity. There is convergent validity in terms of correlation with the technology readiness index that addresses discomfort and insecurity about technological developments but discriminant validity in terms of innovativeness. Therefore, the scale covers both stereotypes'

attributes on the perception of augmented humans and technological attributes. In Survey #3, we could also show concurrent validity in that attitudes toward augmented humans can predict whether participants are willing to use augmentation technologies themselves. Therefore, positive attitudes regarding threat and control in augmented humans are associated with acceptance of the technology. This mirrors the recent call in HCI [43] to integrate negative aspects of social acceptability into technological acceptance models.

So far, research in the area of social attitudes toward augmented humans has been limited due to the lack of assessment tools. Work that considered attitudes towards augmented humans was mainly conducted using qualitative methods [77, 82]. Quantitative studies in the domain have adapted conventional scales ,e.g. from the SCM [13, 52, 56], at the expense of interoperability and specificity to the domain of human augmentation. SHAPE now gives researchers in the domain of human augmentation a tool to quantify attitudes in terms of Social Threat and Agency, which adds a quantitative tool to the repertoire of researchers in the domain of human augmentation. We envision that the scale can meaningfully complement qualitative approaches and thereby enable holistic and impactful insights into the field of human augmentations. In this respect, the scale can be a particularly valuable addition when it comes to comparative long-term studies and studies that are concerned with the attitudes of different samples towards human augmentation technologies (e.g., users from different countries).

According to Villa et al. [82], new augmentation devices should be designed with a focus not only on the artifact itself but on the human that would be integrating it into their life/body and their social environment. Our scale development process showed that the assessment of the social human factor is comprised of two aspects: Social Threat and Agency, which should be considered when evaluating ATs and other types of performance-enhancing technologies.

In the final set of items of SHAPE , there is no explicit reference to privacy threats, which is interesting given that only one item related to privacy was removed, while the remaining items underwent filtering in the EFA. The absence of explicit representation of privacy concerns among the filtered items may suggest that we considered them to be less relevant compared to other factors, such as the agency of the augmented human or the perceived threat it poses to the observer. Furthermore, we acknowledge that the subscale "Social Threat" may not specifically target any particular type of threat, including privacy threats. Therefore, it is possible that this subscale captures certain aspects of privacy concerns, even in the absence of explicit references to privacy threats.

Our study provides valuable information on the social perception of augmented humans. In the initial item pool, we had a sizable number of items that corresponded to a benevolent or positive view of augmented humans,e.g., "An augmented human is interesting." or "An augmented human is friendly." from the MAS scale. However, none of these items surfaced in the exploratory factor analysis to correspond to a factor. We thus suggest that in our sample, attitudes mainly revolved around a negative view of augmented humans. This aligns with recent scale developments such as the Creepiness of Technology Scale (PCTS) [84] where the authors reported three subscales, all of them negatively valenced. Nevertheless, it will be important for future research to investigate measurement invariance of the SHAPE scale as attitudes differ across cultures [82]. This resonates with the fact that beliefs and attitudes toward innovative technologies are ever-evolving. To illustrate this point, the TRI was updated after only a little more than a decade [62, 63] to cover novel aspects of technology readiness. Likely SHAPE might need to be revised when augmentation technologies are more broadly used. This limitation also points to the research opportunity to investigate with the SHAPE scale how attitudes evolve and change over time. The SHAPE scale was built to be unspecific concerning the disability status and the type of augmentation, covering sensory, motor and cognitive augmentations alike; future studies may piece apart how attitudes differ as a function of augmentation characteristics and person characteristics. In order to enable this, we validated the disabilities scenario and discovered that SHAPE can also be utilized effectively in the context of disabilities by ignoring the non-descriptive items. The three non-descriptive items for the disability case pertain to situations that may have been affected by the observer's forgiveness of individuals with prior disabilities. This aligns with

previous work that has found that observers find more acceptable the use of some technologies when the user has a disability condition [23, 34, 66]. In subsection 7.1 we provide a tailored scoring system for this specific case.

The final version of the SHAPE scale is available at Anonymized.com. This website has long-term support planned and is available to distribute the scale easily. The website is planned to serve as a reference point to evaluate the evolution of the attitudes toward human augmentation and performance-enhancement technologies. The anonymized collected data in the website along with translated versions of the SHAPE scale will be made available for researchers to further advance the field.

7.1 Scoring

The SHAPE scale is scored on a seven-point Likert scale from Not at All (1) to Very Much (7). Items S4, S7 and S13 are reverse-scored. Higher scores indicate higher aversion towards AT's users:

7.1.1 Full Scoring System. In the full scoring system of the SHAPE scale, it is advisable to calculate the arithmetic mean of all the items to obtain the overall score, or to compute the mean of the items corresponding to each subscale if the reader seeks insights into specific dimensions. This approach is feasible because both subscales possess equal valence; higher scores indicate a greater degree of aversion towards Augmented Humans or Performance Enhancing technology users.

$$SHAPE = \frac{\mu_{ST} + \mu_{AG}}{2}$$
$$with\ \mu_{ST} = mean(S1 + S2 + S3 + S4_R + S5 + S6 + S7_R),$$
$$and\ \mu_{AG} = mean(S8 + S9 + S10 + S11 + S12 + S13_R)$$

7.1.2 Disability Scenarios Scoring System. In the context of disability scenarios, we suggest using a scoring system similar to the full scoring system. However, instead of utilizing the entire set of items, we suggest excluding the items that exhibit a pronounced skew in opinions toward individuals with disabilities. This adjustment is intended to improve the reliability and validity of the scoring procedure.

$$SHAPE = \frac{\mu_{ST} + \mu_{AG}}{2}$$
$$with\ \mu_{ST} = mean(S1 + S3 + S4_R + S5 + S6 + S7_R),$$
$$and\ \mu_{AG} = mean(S8 + S9 + S10 + S13_R)$$

8 CONCLUSION

We present a measure for assessing attitudes toward augmented humans. The SHAPE scale presented high internal consistency, reliability and high Concurrent, Convergent and Divergent validity. The SHAPE scale is a useful tool to systematically investigate and analyze social attitudes during the design of digital technologies that aim to enhance human performance.

The SHAPE scale also aims to facilitate the integration of a user-centered approach in this field, which was previously characterized by focusing on technical developments and exploratory qualitative methods. With this, the ultimate goal is to enable the development of functional human augmentation technologies that meet the needs and preferences in terms of the user and their social environment. Importantly, it is crucial to emphasize that the development of the SHAPE scale does not intend to assert a hierarchical judgment regarding the merits of qualitative versus quantitative data collection methodologies, but rather aims to enhance the methodological comprehensiveness and rigor of AHs research within the HCI domain.

The SHAPE scale is a suitable and brief instrument for studying the evolution of attitudes toward human augmentation over time. In order to facilitate this goal, we also developed a long-term support website (shape. medien.ifi.lmu.de) to sample these attitudes over time and enable academics and designers to better understand the societal dynamics surrounding ATs.

9 OPEN SCIENCE

This research adheres to the principles of open science and research transparency; The data and materials used in this study, including the protocol, survey data, and statistical code, are publicly available at https://osf.io/9fjvh/. We believe that it is important that other researchers are able to verify and expand our findings, and we encourage others to use our data and methods in their studies.

ACKNOWLEDGMENTS

REFERENCES

[1] Yomna Abdelrahman, Pascal Knierim, Pawel W Wozniak, Niels Henze, and Albrecht Schmidt. 2017. See through the fire: evaluating the augmentation of visual perception of firefighters using depth and thermal cameras. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers.* ACM International Joint Conference on Pervasive and Ubiquitous Computing, 693–696.

[2] Yomna Abdelrahman, Alireza Sahami Shirazi, Niels Henze, and Albrecht Schmidt. 2015. Investigation of material properties for thermal imaging-based interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems.* ACM Conference on Human Factors in Computing Systems, 15–18.

[3] Yomna Abdelrahman and Albrecht Schmidt. 2018. Beyond the visible: sensing with thermal imaging. *Interactions* 26, 1 (2018), 76–78.

[4] Ananta Narayanan Balaji and Li-Shiuan Peh. 2021. AI-on-Skin: Enabling On-Body AI Inference for Wearable Artificial Skin Interfaces. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(Chi Ea '21).* Association for Computing Machinery, New York, NY, USA, Article 358, 7 pages. https://doi.org/10.1145/3411763.3451689

[5] Owen N. Beck, Paolo Taboga, and Alena M. Grabowski. 2022. Sprinting with prosthetic versus biological legs: insight from experimental data. *Royal Society Open Science* 9 (2022).

[6] Dan Bennett, Oussama Metatla, Anne Roudaut, and Elisa D. Mekler. 2023. How Does HCI Understand Human Agency and Autonomy?. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23).* Association for Computing Machinery, New York, NY, USA, Article 375, 18 pages. https://doi.org/10.1145/3544548.3580651

[7] Marit Bentvelzen, Jasmin Niess, Mikołaj P. Woźniak, and Paweł W. Woźniak. 2021. The Development and Validation of the Technology-Supported Reflection Inventory. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(Chi '21).* Association for Computing Machinery, New York, NY, USA, Article 366, 8 pages. https://doi.org/10.1145/3411764.3445673

[8] Hardik Bhimani, Anne-Laure Mention, and Pierre-Jean Barlatier. 2019. Social media and innovation: A systematic literature review and future research directions. *Technological Forecasting and Social Change* (2019).

[9] J Martin Bland and Douglas G. Altman. 1999. Measuring agreement in method comparison studies. *Statistical Methods in Medical Research* 8 (1999), 135–160.

[10] Godfred O. Boateng, Torsten B. Neilands, Edward A. Frongillo, Hugo R. Melgar-Quiñonez, and Sera L. Young. 2018. Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health* 6 (2018). https://doi.org/10.3389/fpubh.2018.00149

[11] Emory S Bogardus. 1933. A social distance scale. *Sociology & Social Research* (1933).

[12] Gunnar Breivik, Dag Vidar Hanstad, and Sigmund Loland. 2009. Attitudes towards use of performance-enhancing substances and body modification techniques. A comparison between elite athletes and the general population. *Sport in Society* 12, 6 (2009), 737–754. https://doi.org/10.1080/17430430902944183 arXiv:https://doi.org/10.1080/17430430902944183

[13] Maximilian Bretschneider, Sarah Mandl, Anja Strobel, Frank Asbrock, and Bertolt Meyer. 2022. Social perception of embodied digital technologies—a closer look at bionics and social robotics. *Gruppe. Interaktion. Organisation. Zeitschrift Für Angewandte Organisationspsychologie (GIO)* 53, 3 (2022), 343–358.

[14] M. Brown, N. Tsagarakis, and D.G. Caldwell. 2003. Exoskeletons for human force augmentation. *Industrial Robot: An International Journal* 30, 6 (Dec. 2003), 592–602.

[15] Heng Cao, Zhengyang Ling, Jun Zhu, Yu Wang, and Wei Wang. 2009. Design frame of a leg exoskeleton for load-carrying augmentation. Ieee, 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO).

[16] Raymond Bernard Cattell and Silke Vogelmann. 1977. A Comprehensive Trial Of The Scree And Kg Criteria For Determining The Number Of Factors. *Multivariate behavioral research* 12 3 (1977), 289–325.

[17] Andrew L. Comrey. 1988. Factor-analytic methods of scale development in personality and clinical psychology. *Journal of consulting and clinical psychology* 56 5 (1988), 754–61.

[18] Erin C. Conrad, Stacey Humphries, and Anjan Chatterjee. 2019. Attitudes Toward Cognitive Enhancement: The Role of Metaphor and Context. *AJOB Neuroscience* 10, 1 (2019), 35–47. https://doi.org/10.1080/21507740.2019.1595771 arXiv:https://doi.org/10.1080/21507740.2019.1595771 Pmid: 31070552.

[19] Patrick W Corrigan, Annette Backs Edwards, Amy Green, Sarah Lickey Diwan, and David L Penn. 2001. Prejudice, social distance, and familiarity with mental illness. *Schizophrenia bulletin* 27, 2 (2001), 219–225.

[20] Jose M Cortina. 1993. What is coefficient alpha? An examination of theory and applications. *Journal of applied psychology* 78, 1 (1993), 98.

[21] Yuri P. Danilov, Mitchell E. Tyler, and Kurt A. Kaczmarek. 2008. Vestibular sensory substitution using tongue electrotactile display. Birkhäuser Basel.

[22] F. A. J. de Laat, Bart van Heerebeek, and Jaap J. van Netten. 2019. Advantages and disadvantages of interdisciplinary consultation in the prescription of assistive technologies for mobility limitations. *Disability and Rehabilitation: Assistive Technology* 14 (2019), 386–390.

[23] Richard W. DeVaul, Alex "Sandy" Pentland, and Vicka R. Corey. 2003. The Memory Glasses: Subliminal vs. Overt Memory Support with Imperfect Information. In *Proceedings of the 7th IEEE International Symposium on Wearable Computers (ISWC '03)*. IEEE Computer Society, USA, 146.

[24] Anne M. Dijkstra and Mirjam Schuijff. 2016. Public opinions about human enhancement can enhance the expert-only debate: A review study. *Public Understanding of Science* 25 (2016), 588–602.

[25] Tiffany D. Do, Ryan P. McMahan, and Pamela J. Wisniewski. 2022. A New Uncanny Valley? The Effects of Speech Fidelity and Human Listener Gender on Social Perceptions of a Virtual-Human Speaker. 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 424, 11 pages. https://doi.org/10.1145/3491102.3517564

[26] Shirley A. Elprama, Bram Vanderborght, and An Jacobs. 2021. An industrial exoskeleton user acceptance framework based on a literature review of empirical studies. *Applied ergonomics* 100 (2021), 103615.

[27] Liora Findler, Noa Vilchinsky, and Shirli Werner. 2007. The Multidimensional Attitudes Scale Toward Persons With Disabilities (MAS): Construction and Validation. *Rehabilitation Counseling Bulletin* 50, 3 (2007), 166–176. https://doi.org/10.1177/00343552070500030401 arXiv:https://doi.org/10.1177/00343552070500030401

[28] Susan T. Fiske, Amy J. C. Cuddy, and Peter Glick. 2007. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences* 11, 2 (2007). https://doi.org/10.1016/j.tics.2006.11.005

[29] Susan T. Fiske, Amy J. C. Cuddy, Peter Glick, and Jun Xu. 2002. A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology* 82, 6 (2002). https://doi.org/10.1037/0022-3514.82.6.878

[30] Dirk Folkerts, Roland Loh, Andrea Petróczi, and Sebastian Brueckner. 2021. The performance enhancement attitude scale (PEAS) reached 'adulthood': Lessons and recommendations from a systematic review and meta-analysis. *Psychology of Sport and Exercise* 56 (2021), 101999. https://doi.org/10.1016/j.psychsport.2021.101999

[31] Jesse Graham, Brian A Nosek, Jonathan Haidt, Ravi Iyer, Koleva Spassena, and Peter H Ditto. 2008. Moral foundations questionnaire. *Journal of Personality and Social Psychology* (2008).

[32] Graciela Guerrero, Fernando José Mateus da Silva, Antonio Fernández-Caballero, and António Pereira. 2022. Augmented Humanity: A Systematic Mapping Review. *Sensors* 22, 2 (Jan. 2022), 514. https://doi.org/10.3390/s22020514

[33] Luke Haliburton, Sinksar Ghebremedhin, Robin Welsch, Albrecht Schmidt, and Sven Mayer. 2023. Investigating Labeler Bias in Face Annotation for Machine Learning. https://doi.org/10.48550/arXiv.2301.09902 arXiv:2301.09902 [cs].

[34] Steve Hodges, Emma Berry, and Ken Wood. 2011. SenseCam: A wearable camera that stimulates and rehabilitates autobiographical memory. *Memory* 19, 7 (2011), 685–696. https://doi.org/10.1080/09658211.2011.605591 arXiv:https://doi.org/10.1080/09658211.2011.605591 PMID: 21995708.

[35] John Louis Horn. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30 (1965), 179–185.

[36] Aike C. Horstmann and Nicole C. Krämer. 2020. Expectations vs. actual behavior of a social robot: An experimental investigation of the effects of a social robot's interaction skill level and its expected future role on people's evaluations. *Plos One* 15, 8 (08 2020), 1–18. https://doi.org/10.1371/journal.pone.0238133

[37] Jochen Huber, Roy Shilkrot, Pattie Maes, and Suranga Nanayakkara (Eds.). 2018. *Assistive Augmentation.* Springer Singapore.

[38] Masahiko Inami, Daisuke Uriu, Zendai Kashino, Shigeo Yoshida, Hiroto Saito, Azumi Maekawa, and Michiteru Kitazaki. 2022. Cyborgs, Human Augmentation, Cybernetics, and JIZAI Body. In *Augmented Humans 2022* (Kashiwa, Chiba, Japan) *(AHs 2022)*. Association for

Computing Machinery, New York, NY, USA, 230–242. https://doi.org/10.1145/3519391.3519401

[39] Shunichi Kasahara, Jun Nishida, and Pedro Lopes. 2019. Preemptive Action: Accelerating Human Reaction Using Electrical Muscle Stimulation Without Compromising Agency. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(Chi '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300873

[40] Norene Kelly and Stephen B Gilbert. 2016. The WEAR Scale: Developing a Measure of the Social Acceptability of a Wearable Device. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (2016).

[41] Rafal Kocielnik, Saleema Amershi, and Paul N. Bennett. 2019. Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-User Expectations of AI Systems. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(Chi '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300641

[42] 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* (Honolulu, HI, USA) *(Chi '20)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3313831.3376162

[43] 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* (Honolulu, HI, USA) *(Chi '20)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3313831.3376162

[44] 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

[45] Amel Ksibi, Ala Saleh D. Alluhaidan, Amina Salhi, and Sahar A. El-Rahman. 2021. Overview of Lifelogging: Current Challenges and Advances. *IEEE Access* 9, 62630–62641.

[46] Vikash Kumar, Yogesh Vijay Hote, and Shivam Jain. 2019. Review of Exoskeleton: History, Design and Control. *2019 3rd International Conference on Recent Developments in Control, Automation & Power Engineering (RDCAPE)* (2019), 677–682.

[47] Shun Yin Lam, Jeongwen Chiang, and A. Parasuraman. 2008. The effects of the dimensions of technology readiness on technology acceptance: An empirical analysis. *Journal of Interactive Marketing* 22, 4 (2008), 19–39. https://doi.org/10.1002/dir.20119

[48] Florian Lang, Albrecht Schmidt, and Tonja Machulla. 2020. Augmented Reality for People with Low Vision: Symbolic and Alphanumeric Representation of Information. *Computers Helping People with Special Needs* 12376 (2020), 146–156.

[49] Franklin Mingzhe Li, Di Laura Chen, Mingming Fan, and Khai N. Truong. 2021. "I Choose Assistive Devices That Save My Face": A Study on Perceptions of Accessibility and Assistive Technology Use Conducted in China. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021).

[50] Robert C. MacCallum, Mary Ann Roznowski, and Lawrence B. Necowitz. 1992. Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychological bulletin* 111 3 (1992), 490–504.

[51] Azumi Maekawa, Shota Takahashi, MHD Yamen Saraiji, Sohei Wakisaka, Hiroyasu Iwata, and Masahiko Inami. 2019. Naviarm: Augmenting the Learning of Motor Skills Using a Backpack-Type Robotic Arm System. In *Proceedings of the 10th Augmented Human International Conference 2019* (Reims, France) *(Ah2019)*. Association for Computing Machinery, New York, NY, USA, Article 38, 8 pages. https://doi.org/10.1145/3311823.3311849

[52] Sarah Mandl, Maximilian Bretschneider, Frank Asbrock, Bertolt Meyer, and Anja Strobel. 2022. The Social Perception of Robots Scale (SPRS): Developing and Testing a Scale for Successful Interaction Between Humans and Robots. In *Collaborative Networks in Digitalization and Society 5.0: 23rd IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2022, Lisbon, Portugal, September 19–21, 2022, Proceedings*. Springer, 321–334.

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

[54] Simone Marcheschi, Fabio Salsedo, Marco Fontana, and Massimo Bergamasco. 2011. Body Extender: Whole body exoskeleton for human power augmentation. In *2011 IEEE International Conference on Robotics and Automation*. Ieee, 2011 IEEE International Conference on Robotics and Automation.

[55] Kevin McKee, Xuechunzi Bai, and Susan Fiske. 2021. Understanding human impressions of artificial intelligence. (2021).

[56] Bertolt Meyer and Frank Asbrock. 2018. Disabled or cyborg? How bionics affect stereotypes toward people with physical disabilities. *Frontiers in psychology* 9 (2018), 2251.

[57] James W. Moore. 2016. What Is the Sense of Agency and Why Does it Matter? *Frontiers in Psychology* 7 (2016).

[58] Florian Floyd Mueller, Pedro Lopes, Paul Strohmeier, Wendy Ju, Caitlyn Seim, Martin Weigel, Suranga Nanayakkara, Marianna Obrist, Zhuying Li, Joseph Delfa, Jun Nishida, Elizabeth M. Gerber, Dag Svanaes, Jonathan Grudin, Stefan Greuter, Kai Kunze, Thomas Erickson, Steven Greenspan, Masahiko Inami, Joe Marshall, Harald Reiterer, Katrin Wolf, Jochen Meyer, Thecla Schiphorst, Dakuo Wang, and Pattie Maes. 2020. Next Steps for Human-Computer Integration. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(Chi '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376242

[59] Junichi Nabeshima, MHD Yamen Saraiji, and Kouta Minamizawa. 2019. Prosthetic Tail: Artificial Anthropomorphic Tail for Extending Innate Body Functions. In *Proceedings of the 10th Augmented Human International Conference 2019* (Reims, France) *(Ah2019)*. Association for Computing Machinery, New York, NY, USA, Article 36, 4 pages. https://doi.org/10.1145/3311823.3311848

[60] Mine Oyman, Dondu Bal, and Serhat Ozer. 2022. Extending the technology acceptance model to explain how perceived augmented reality affects consumers' perceptions. *Computers in Human Behavior* 128 (2022), 107127. https://doi.org/10.1016/j.chb.2021.107127

[61] Claudio Pacchierotti, Stephen Sinclair, Massimiliano Solazzi, Antonio Frisoli, Vincent Hayward, and Domenico Prattichizzo. 2017. Wearable Haptic Systems for the Fingertip and the Hand: Taxonomy, Review, and Perspectives. *EEE Trans. Haptics* 10, 4 (Oct. 2017), 580–600. https://doi.org/10.1109/toh.2017.2689006

[62] A. Parasuraman. 2000. Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. *Journal of Service Research* 2, 4 (2000), 307–320. https://doi.org/10.1177/109467050024001 arXiv:https://doi.org/10.1177/109467050024001

[63] A. Parasuraman and Charles L. Colby. 2015. An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research* 18, 1 (2015), 59–74. https://doi.org/10.1177/1094670514539730 arXiv:https://doi.org/10.1177/1094670514539730

[64] Andrea Petróczi and Eugene Aidman. 2009. Measuring explicit attitude toward doping: Review of the psychometric properties of the Performance Enhancement Attitude Scale. *Psychology of Sport and Exercise* 10, 3 (2009), 390–396. https://doi.org/10.1016/j.psychsport.2008.11.001

[65] Domenico Prattichizzo, Maria Pozzi, Tommaso Lisini Baldi, Monica Malvezzi, Irfan Hussain, Simone Rossi, and Gionata Salvietti. 2021. Human augmentation by wearable supernumerary robotic limbs: review and perspectives. *Progress in Biomedical Engineering* 3, 4 (Sept. 2021), 042005. https://doi.org/10.1088/2516-1091/ac2294

[66] Halley Profita, Reem Albaghli, Leah Findlater, Paul Jaeger, and Shaun K. Kane. 2016. The AT Effect: How Disability Affects the Perceived Social Acceptability of Head-Mounted Display Use. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 4884–4895. https://doi.org/10.1145/2858036.2858130

[67] Michael J. Proulx. 2010. Synthetic synaesthesia and sensory substitution. *Consciousness and Cognition* 19 (2010), 501–503.

[68] Roope Raisamo, Ismo Rakkolainen, Päivi Majaranta, Katri Salminen, Jussi Rantala, and Ahmed Farooq. 2019. Human augmentation: Past, present and future. *International Journal of Human-Computer Studies* 131 (2019), 131–143.

[69] Jie Ren, Ling Gao, Xiaoming Wang, Miao Ma, Guoyong Qiu, Hai Wang, Jie Zheng, and Zheng Wang. 2022. Adaptive Computation Offloading for Mobile Augmented Reality. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 175 (Dec. 2022), 30 pages. https://doi.org/10.1145/3494958

[70] Rebekah A. Rousi and Roni Renko. 2020. Emotions toward cognitive enhancement technologies and the body - Attitudes and willingness to use. *Int. J. Hum. Comput. Stud.* 143 (2020), 102472.

[71] Tomoya Sasaki, MHD Yamen Saraiji, Kouta Minamizawa, and Masahiko Inami. 2018. MetaArms: Body Remapping Using Feet-Controlled Artificial Arms. In *Adjunct Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 140–142. https://doi.org/10.1145/3266037.3271628

[72] Astrid Schepman and Paul Rodway. 2020. Initial validation of the general attitudes towards Artificial Intelligence Scale. *Computers in Human Behavior Reports* 1 (2020), 100014. https://doi.org/10.1016/j.chbr.2020.100014

[73] Albrecht Schmidt. 2017. Augmenting human intellect and amplifying perception and cognition. *IEEE Pervasive Computing* 16, 1 (2017), 6–10.

[74] Ali Shafti, Shlomi Haar, Renato Mio, Pierre Guilleminot, and A. Aldo Faisal. 2021. Playing the piano with a robotic third thumb: assessing constraints of human augmentation. *Scientific Reports* 11, 1 (Nov. 2021), 21375. https://doi.org/10.1038/s41598-021-00376-6

[75] Kristen Shinohara and Jacob O. Wobbrock. 2011. In the shadow of misperception: assistive technology use and social interactions. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011).

[76] Ben Shneiderman. 2020. Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction* 36, 6 (2020), 495–504. https://doi.org/10.1080/10447318.2020.1741118 arXiv:https://doi.org/10.1080/10447318.2020.1741118

[77] Sandra Maria Siedl and Martina Mara. 2022. What Drives Acceptance of Occupational Exoskeletons? Focus Group Insights from Workers in Food Retail and Corporate Logistics. *International Journal of Human-Computer Interaction* (2022), 1–10.

[78] Taufik Akbar Sitompul and Markus Wallmyr. 2019. Using Augmented Reality to Improve Productivity and Safety for Heavy Machinery Operators: State of the Art. In *Proceedings of the 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Brisbane, QLD, Australia) *(Vrcai '19)*. Association for Computing Machinery, New York, NY, USA, Article 8, 9 pages. https://doi.org/10.1145/3359997.3365689

[79] David L. Streiner. 1994. Figuring Out Factors: The Use and Misuse of Factor Analysis. *The Canadian Journal of Psychiatry* 39 (1994), 135–140.

[80] Hamed Taherdoost. 2018. A review of technology acceptance and adoption models and theories. *Procedia Manufacturing* 22 (2018), 960–967.

[81] 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). https://doi.org/10.3389/fpsyg.2017.01552

[82] Steeven Villa, Jasmin Niess, Takuro Nakao, Jonathan Lazar, Albrecht Schmidt, and Tonja-Katrin Machulla. 2023. Understanding Perception of Human Augmentation: A Mixed-Method Study. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany* (April 2023). https://doi.org/10.1145/3544548.3581485

[83] Robin Welsch, Alexander F. Schmidt, Daniel Turner, and Martin Rettenberger. 2021. Test–Retest Reliability and Temporal Agreement of Direct and Indirect Sexual Interest Measures. *Sexual Abuse* 33, 3 (2021), 339–360. https://doi.org/10.1177/1079063220904354 arXiv:https://doi.org/10.1177/1079063220904354 Pmid: 32036763.

[84] Paweł W. Woźniak, Jakob Karolus, Florian Lang, Caroline Eckerth, Johannes Schöning, Yvonne Rogers, and Jasmin Niess. 2021. Creepy Technology:What Is It and How Do You Measure It?. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(Chi '21)*. Association for Computing Machinery, New York, NY, USA, Article 719, 13 pages. https://doi.org/10.1145/3411764.3445299

[85] Di Wu, Jinhui Ouyang, Ningyi Dai, Mingzhu Wu, Haodan Tan, Hanhui Deng, Yongmei Fan, Dakuo Wang, and Zhanpeng Jin. 2022. DeepBrain: Enabling Fine-Grained Brain-Robot Interaction through Human-Centered Learning of Coarse EEG Signals from Low-Cost Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6, 3, Article 147 (Sept. 2022), 27 pages. https://doi.org/10.1145/3550334

[86] Emiliano Quinones Yumbla, Zhi Qiao, Weijia Tao, and Wenlong Zhang. 2021. Human Assistance and Augmentation with Wearable Soft Robotics: a Literature Review and Perspectives. *Current Robotics Reports* 2, 4 (Dec. 2021), 399–413. https://doi.org/10.1007/s43154-021-00067-0

[87] Guangjian Zhang and Kristopher J Preacher. 2015. Factor Rotation and Standard Errors in Exploratory Factor Analysis. *Journal of Educational and Behavioral Statistics* 40 (2015), 579–603.

A THE SHAPE SCALE AND SCORING SYSTEM

Table 7. The final version of the SHAPE scale consisting of thirteen items. Internal consistency Cronbach's alpha values are displayed on top of their respective item group. Each item is answered in a 7 point Likert scale ranging from (1) Not at all to (7) Very Much. (*) denotes that the item is inverted.

Item	ID	Source
Social Threat $\alpha = 0.808$		
An augmented human is a threat to society.	**S1**	CS
An augmented human would be dangerous.	**S2**	CS
An augmented human is intimidating.	**S3**	MAS
(*) An augmented human would conform to the traditions of society.	**S4**	MFQ
An augmented human has to disclose their augmentation.	**S5**	CS
An augmented human would do something cruel	**S6**	MFQ
(*) An augmented human is more competitive than a non-augmented human.	**S7**	PeaS
Agency $\alpha = 0.809$		
The actions of the augmented human do not match their intentions.	**S8**	SoA
An augmented human is not the author of their own actions.	**S9**	SoA
An augmented human is just an instrument of something or somebody else.	**S10**	SoA
An augmented human does things without any intention.	**S11**	SoA
An augmented human suffering through their augmentation should get help.	**S12**	MFQ
(*)An augmented human is in full control of what they do.	**S13**	SoA

A.1 Full scoring system

In the full scoring system of the SHAPE scale, it is advisable to calculate the arithmetic mean of all the items to obtain the overall score, or to compute the mean of the items corresponding to each subscale if the reader seeks insights into specific dimensions. This approach is feasible because both subscales possess equal valence; **higher scores indicate a greater degree of aversion towards Augmented Humans or Performance Enhancing technology users.**

$$SHAPE = \frac{\mu_{ST} + \mu_{AG}}{2}$$

$$with\ (SocialThreat\ Subscale)\mu_{ST} = mean(S1 + S2 + S3 + S4_R + S5 + S6 + S7_R),$$

$$and\ (Agency\ Subscale)\mu_{AG} = mean(S8 + S9 + S10 + S11 + S12 + S13_R)$$

A.2 Disability Scenarios Scoring System

In the context of disability scenarios, we suggest using a scoring system similar to the full scoring system. However, instead of utilizing the entire set of items, we suggest excluding the items that exhibit a pronounced skew in opinions toward individuals with disabilities. This adjustment is intended to improve the reliability and validity of the scoring procedure.

$$SHAPE = \frac{\mu_{ST} + \mu_{AG}}{2}$$

$$with\ (SocialThreat\ Subscale)\mu_{ST} = mean(S1 + S3 + S4_R + S5 + S6 + S7_R),$$

$$and\ (Agency\ Subscale)\mu_{AG} = mean(S8 + S9 + S10 + S13_R)$$