A Nine-Item Questionnaire for Measuring the Social Disfordance of Mediated Social Touch Technologies

KENYA MEJIA, University of Minnesota
SVETLANA YAROSH, University of Minnesota[1]

Mediated Social Touch (MST) technologies focus on enhancing a communication experience by sensing, transmitting, and simulating social touch between remote partners. With interest in developing MST technologies continuing to grow, it is important to create standardized methods for measuring the effect of these novel systems. We designed and validated a 9-item questionnaire to measure the "Social Disfordance" of Mediated Social Touch, with three scales that focus on Social Discomfort, Communicational Expressiveness, and Need for Additional Consideration. A high degree of "social disfordance" of an MST system signifies that it may not provide the appropriate social affordances for mediating touch in a particular context. The development of the Social Disfordance of Mediated Social Touch (SDMST) instrument included a systematic literature review, expert feedback, and think-out-loud piloting. Its refinement included an exploratory factor analysis with a subsequent reduction of questions and scales. We report its psychometric properties, including metrics of inter-item reliability, convergent validity, test-retest reliability, and concurrent validity, confirming that these properties are sufficient for future use. We conclude with examples of scoring, appropriate use, and a discussion of the limitations.

CCS Concepts: • **Human Computer Interaction (HCI)** → Collaborative and social computing→ Collaborative and social computing design and evaluation methods

KEYWORDS: Mediated social touch; haptics; computer-mediated communication; validated questionnaire; social disfordance

ACM Reference format:

Kenya Mejia and Svetlana Yarosh. 2018. A Nine-Item Questionnaire for Measuring the Social Disfordance of Mediated Social Touch Technologies. PACM on Human-Computer Interaction, 1, CSCW, Article 77 (November 2017), 17 pages. https://doi.org/10.1145/3134712

1 INTRODUCTION

Touch is the first sense human beings develop and it is key to our relationships with our family, friends, and other members of our communities [16]. Humans build and maintain social relationships through touch by using it to communicate friendliness, affection, intimacy, support, playfulness, and more [25]. Just as video-mediated communication technologies, like video chat, allow users to stay in touch with friends and family through video and audio, mediated social touch (abbreviated as MST) is defined as technologies that allow

[1] This.is work was in part supported by the National Science Foundation, under grant CHS-1526085.
Corresponding author's address: Svetlana Yarosh, Keller Hall, 200 Union St. SE Room 4-192, Minneapolis, MN 55455-0159.

people to touch across distance in order to reinforce social relationships [4]. However, unlike video and audio, mediated transmission of touch cannot occur at nearly the same fidelity because it is such an embodied action. Human perceptions of social touch reflect thousands of years of evolutionary development. Specific types of touch are suitable or unsuitable for specific relationships, contexts, and tasks [25]. This constitutes a challenging design space.

As interest in communication technologies for enhancing social relationships grows, so does the development of new MST systems. Given that novel MST systems tend to not be robust enough for field deployments, most of these systems are evaluated in controlled or lab settings (e.g., [1,7,25,33,40]). Previous evaluations have focused on qualitative reflections or responses to questions written by the investigator for the specific purpose of that study. This makes it difficult to compare across investigations and across design alternatives [18]. In particular, prior to this work, there has been no standardized way to assess how a particular MST technology may support or hinder a social communication experiences. In this work, we develop and validate a questionnaire instrument for this purpose; in order to provide a quantitative complement to qualitative methods to continue developing systematic research methodologies for this embryonic field [15].

We use the term "social disfordance" (coined as the opposite of social affordance [10]) to refer to cases where an MST technology may hinder social connection in a specific contact. The higher the social disfordance, the more the technology hindered the social interactions. The initial development of the Social Disfordance of Mediated Social Touch (SDMST) questionnaire included a literature review, expert feedback, and think-out-loud piloting. Its refinement included an exploratory factor analysis with a subsequent reduction of questions and scales, resulting in a 3-scale, 9-item questionnaire. Finally, various metrics for validation confirmed that the SDMST questionnaire's psychometric properties are sufficient for use.

We begin by discussing related work associated with mediated social touch and questionnaire validation in HCI and related fields. Next, we describe the process of designing and developing a questionnaire, starting with a systematic literature review and then reporting how expert and pilot participant feedback was used to iterate and refine it. We discuss the use of an exploratory factor analysis to reduce and optimize the number of questions. We report the questionnaire's validation, which includes inter-item reliability, convergent validity, test-retest reliability, and concurrent validity. Finally, the paper's discussion provides guidance on questionnaire scoring, appropriate use, and reflections on the limitations of the SDMST.

2 RELATED WORK

Haptic communication technologies and Mediated Social Touch (MST) have been of significant interest to the CSCW community (e.g., [2,8,32,42,52]). In this section, we discuss the major threads of research in both of these approaches, highlighting the difference between haptic communication and MST. We also discuss the previous work done to measure the effects of MST technologies.

2.1 Haptic Technologies for Communication and Mediated Social Touch

Our review found that previous work on haptic technologies for communication falls roughly into three major categories: haptic awareness mechanisms, haptic feedback in collaboration, and haptics as a channel for information signal. The first two approaches focus on offloading information onto the haptic channel as a strategy for managing information overload and increasing the amount of feedback that can be meaningfully processed by the user, such as an alternative channel for notifications and awareness (e.g., [40]). The third approach focuses on the use of the sense of touch as a channel to transmit information and investigate human ability to interpret such signal. For example, several studies asked participants to transmit Morse code messages to each other using vibrotactile or pressure-based hand devices (e.g., [7,29]). Other studies focus on developing and validating alternative alphabets for information transmission by developing units of meaning known as tactons [5], tactile icons [23], or haptic phonemes [14]. Another

significant class of devices focuses on "affective haptics" which attempt to represent and transmit specific emotions as tactile sensations (reviewed in [13]).

In contrast, Mediated Social Touch concerns itself not with the transmission of information signal through haptic channels, but rather with reestablishing the social benefits of touch in remote communication contexts. MST is closely related to the concern of "mediating intimate relationships," though not all mediating technology consider the role or importance of touch [22]. One difference between investigations of haptic communication and those of mediated social touch is that haptic technologies, by definition, include some level of physical actuation (e.g., movement, vibration, heat), while mediated social touch technologies may or may not include haptic elements in how they attempt to transmit the experience of social touch. In fact, there are multiple investigation that report that the sensation of touch can be simulated through sensory illusions by taking advantage of the dominance of visual senses to the interpretation of stimuli (e.g., [19,24,30]). One common thread of work in mediated social touch focuses on enhancing existing communication channels by incorporating an expressive haptic channel. These are not meant to replicate an existing touch but rather create a new vocabulary for social interaction. The classic example of this approach is InTouch [3], a networked device that synchronized state across two locations (e.g., pushing a rod locally, moved the rod on a remote device). Similarly, Park et al. [35] incorporated haptic feedback into an off-the-shelf mobile phone system, allowing participants to tap the back of the phone while speaking, vibrating the partner's phone in patterns such as "tickling," "slapping," and "tapping." They deployed it to three couples for a month of use (one of the only examples of a field study in this domain). Another strategy in the Mediated Social Touch domain is in supporting existing forms of social touch, such as shaking hands, hugging, kissing, etc. For example, Nakanishi et al. [32] investigated a robotic arm for transmitting a handshake while videoconferencing and found that lab study participants felt increased social presence compared to the videoconference only condition. Several researchers investigated haptic hugging as a possible interaction, for example Tsetserukou [46] built a vibrotactile vest and Cha et al [6] developed a vibrotactile jacket for transmitting a hug sensation (though these systems were not investigated in user evaluations).

Overall, while researchers have been successful in designing and prototyping new devices for haptic communication and Mediated Social Touch, multiple systematic reviews in this space point out that these prototypes "have not yet been submitted to empirical scrutiny" [18] and largely "suffer from a lack of robust empirical testing" [17]. While these statements were made several years ago, MST work as recent at CSCW 2017 still does not rely on employing validated metrics for evaluation (e.g., [42]). This is not a criticism of the previous research but rather a reflection of the lack of reliable measurement—one of the missing pieces necessary to support advances in the field. We think integrating validated measures could favorably complement the field's current approach of qualitative and design investigations.

2.2 Measuring User Experiences with Mediated-Communication Technologies

Rigorous empirical work typically requires the ability to reliably measure variables of interest. However, currently both research in haptic communication and research in MST lack validated instruments for measuring user perceptions of a technologically mediated social touch experience.

There are a number of questionnaire instruments in social science that one might consider using to evaluate the role of technologies in social communication. One approach is deploying a technology and collecting pre- and post-deployment metrics to detect changes in relationship quality or mood. For example, a person's general positive and negative affect can be measured using the Positive and Negative Affect Schedule by collecting repeated metrics over a period of several weeks [44]. Similarly, affect could be represented using Russell's circumplex model, which evaluates emotions on the two bipolar dimensions, pleasure-displeasure and arousal-sleepiness [39]. There are also a number of metrics for assessing the quality of any specific relationship, such as the Quality of Relationships Inventory [37]. However, these metrics may not be appropriate for lab-based testing and may not show an effect outside of a long field deployment, so they cannot be used in most investigations in this domain (which occur almost exclusively in the lab at the time of this work).

There have been several attempts to design questionnaire instruments specifically aimed at evaluating communication technologies, some of which can be adapted for the sort of lab study that is typical of work in this domain. Several validated inventories have been developed both in social science (e.g., [27,28]) and in HCI (e.g., [19,50]). Another method adapted for evaluating communication technologies is the layered protocol method, which evaluates the intention to perform an action, limitations of the human body, and specifications of the communication device and the network to derive design guidelines [38]. While some scales on these instruments measure aspects relevant to social touch (e.g., engagement, presence), none of them discuss the sense of touch explicitly and they are each missing core aspects of the MST experience. We found that researchers studying mediated touch most frequently generate their own sets of questions to measure whether the experience was expressive, uncomfortable, distracting, etc. (e.g., [33,49]). Unfortunately, this means that each study asks different questions and these questions are not validated in any way. This approach prevents combining results through systematic meta-analysis, which could allow investigators to draw more confident conclusions across investigations. In this paper, we present the design and validation of the Social Disfordance of Mediated Social Touch (SDMST) metric as a standardized and validated instrument for the evaluation of novel social touch technologies.

3 INITIAL DESIGN

In this section, we describe the initial design and development of the Social Disfordance of Mediated Social Touch (SDMST) questionnaire through a systematic literature review, expert feedback, and think-out-loud piloting. We followed a standard process in questionnaire development and validation, though we combined techniques from multiple previous papers when we thought that the work could be strengthened by it. For example, the process of systematic literature review and think-aloud piloting was done during ABCCT validation [50], while the process of soliciting expert feedback was carried out during the development of the development of the User Burdens scale [43]. We combined these approaches.

3.1 Systematic Literature Review

As in many other areas of user experience, there is a paucity of theory on mediated social touch. Instead of theoretical constructs, we relied on a systematic review of empirical work to help formulate the relevant operationalizations and constructs. We completed a systematic literature review, which included 103 relevant papers on mediated social touch. These papers were collected using Google Scholar as a database. We used incognito mode (to prevent cookies from being locally stored, which would change the order of results) and searched for the phrases "tangible communication," "haptic communication," and "mediated social touch," sorted by relevance. We evaluated the top 100 papers for each search term. The research team read the abstracts of all resulting papers to develop an exclusion criteria. We excluded papers that did not include technological mediation of touch and papers that focused on touch between robots and people (rather than social communication). We applied exclusion criteria to our initial set of 300 papers and removed duplicates. This process yielded 103 total papers. We read the papers and analyzed them according to the technologies, relationships, and situations that each study used to address mediated social touch. More relevant to this investigation, we read through the papers and compiled the questions asked in the various studies. In some cases, the questions were not reported in the paper and we reached out to the authors to get the complete list of questions asked (in all such cases, we were successful in connecting with the authors).

We found that there was quite a bit of diversity in wording and topic among the questions compiled. It was clear that no set of questions could be readily adapted for use with other systems. For example, many questions referred to the specifics of a given task (e.g., handshake) or asked about a specific modality of touch employed in a given system (e.g., heat). In order to abstract to higher-level themes, we followed a standard data-driven thematic analysis approach [41]. We first open-coded the available questions, we clustered the open-codes using affinity mapping, and we labeled the emerging themes as higher-level categories.

Table 1. Questions categories extracted from analysis of previous Mediated Social Touch evaluations

Question Category	Category Definition	Previous Work Example Question
Intensity	Was this interaction intense or calm?	How overwhelmed, or calm, did you feel when being greeted in this manner?
Physical Realism	Was this interaction similar to touching in an unmediated way?	How realistic, or unrealistic, did you find this greeting? (Very unrealistic to Very realistic)
Value Added	Did the touch aspects of the system enhance the communication experience?	How well were you able to express this emotion [with this MST system]?
Social Comfort	Was the way this system transmitted touch natural and socially appropriate for the relationship and context?	What did you perceive as comfortable or uncomfortable during the experiment, if anything?

Through this process, we were able to identify several clusters of questions from previous studies that were not relevant the measurement of MST systems. For example, general demographics, questions regarding task difficulty, questions that referred to specific touch modalities (e.g., "Could you perceive the pattern in the vibration?"), and questions which required the user to express direct preference between a number of alternative systems would not be appropriate to include in a general MST instrument. We also found that a number of categories of questions attempted to measure aspects of systems that already have existing validated metrics. For example, we found that one category of questions related to social presence. We excluded those from subsequent analysis, as there already exist validated metrics for measuring social presence (e.g., Networked Minds Measure of Social Presence [19]). We also excluded the category of questions that asked about the ease-of-use, cognitive load, and other types of workload of the system. Again, this decision was made because there is an existing validated metric that captures these aspects—the NASA-TLX [20].

In the remaining dataset, we found that questions fell into four categories: Intensity, Physical Realism, Value Added, and Social Comfort. Table 1 defines the categories and provides examples of the questions from each of these four major categories. These categories represented the questions that were being asked to

evaluate MST systems, but did not have corresponding standardized wording or measures. In order to provide a questionnaire that would be relevant to the people who would be using it, we contacted experts for feedback on these categories.

3.2 Expert Feedback

For our purpose, we considered any author of a peer-reviewed paper on MST to be a potential expert on the topic. We used our systematic literature review to identify potential experts, emailing the 148 authors of the 103 papers included in our data set (44 of those emails bounced back, mostly from student author emails). Through a Google form distributed over email, authors were invited to provide feedback and discuss the four categories listed in Table 1. We asked them to rate each category's importance to evaluating MST technologies (on a Likert-type scale of importance) and comment as to whether any important aspects of MST systems were missing from the current analysis. We also provided five sample questions for each category to clarify how it may be operationalized in a questionnaire (none of these questions were direct examples from papers, but rather a representative operationalization of each category). We received feedback from 14 authors representing 13.4% of the contacted authors and 18.4% papers (several of them were authors on multiple papers). The respondents were as follows: 9 academic professionals (e.g., professors), 2 industry researchers, 1 post-doctoral researcher, and 2 Ph.D. candidates. We acknowledge that this is a diverse set of experience, which may have influenced their response and perspectives. Figure 1

summarizes the feedback provided by experts. Of the four categories, Social Comfort was the category that every expert thought most relevant to evaluating mediated social touch technologies. Also, in their qualitative feedback, several experts underscored the importance of this scale and encouraged us to pursue it in our investigation. Value Added was the second most highly rated category, though it was less universally deemed as important. The other two categories received less positive and more mixed ratings of importance.

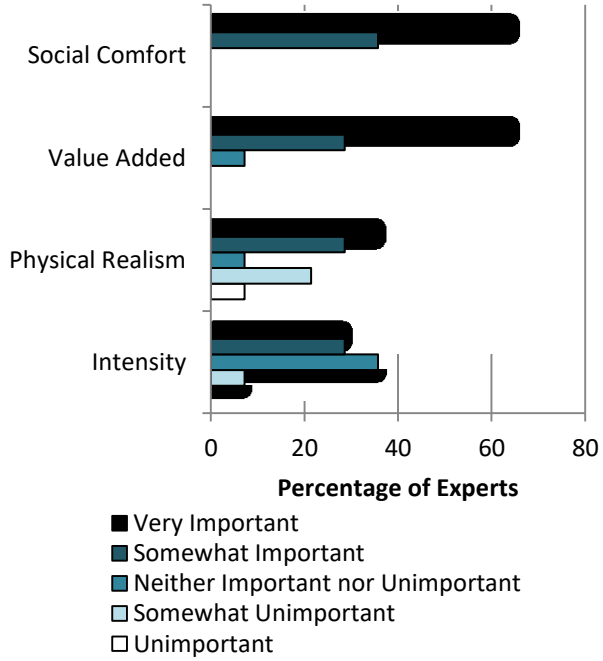

Figure 1. Shows the responses from the experts. The graph shows the categories where experts are in agreement and where there is a range of opinions.

Beyond these specific ratings, in the qualitative feedback, experts expressed their concern about the different factors that could influence the social interaction aside from the technology and encouraged us to target the questions at a specific instance and relationship involved in the interaction. Additionally, several researchers shared other measures that they had found helpful in studying MST technologies (e.g., AttrakDiff [21]) and mentioned that it was important to be able to combine the resulting questionnaire with other methods and metrics. Given the unequivocal importance of the Social Comfort and (to a lesser extent) Value Added categories, we chose to dedicate commensurate focus to these ideas in designing the questionnaire.

3.3 Generating Questions and Piloting

The next step in developing the questionnaire was coming up with questions that would represent the Social Comfort (and to a lesser extent, the Value Added) categories. The research team brainstormed 52 questions that were inspired by expert feedback and that represented alternative aspects and ways of asking about whether an experience with a given MST technology and a given partner was socially natural, beneficial, appropriate, and comfortable.

This set of questions was piloted with three participants through think-out-loud testing. Participants were asked to use an established MST technology with the researcher. The technology used was

"Thumbkiss"—a feature out of the mobile app "Couple."[2] This application delivers tactile feedback when thumbs overlap on the same area of each person's phone screen when two people are using the application for remote communication. Each participant tested out the application and then answered each of the questions on the questionnaire. They were asked to discuss their thought process out loud and generally reflect on anything that was confusing about the question. They were also encouraged to ask clarifying questions. Each participant was audio recorded to document the process and the lead author analyzed these transcripts to identify which questions were most problematic. We ended this phase after three participants because there was substantial agreement among them on which questions were problematic in the set. It was sufficient to support the next iteration of the questionnaire's design.

Through this process, we eliminated three questions and modified six questions to address points of potential confusion. The refined set of 49 questions was used in the next phase.

4 EXPLORATORY FACTOR ANALYSIS AND QUESTIONNAIRE REFINEMENT

In order to continue refining and validating the questionnaire, we deployed it to 114 adult participants, conducted an exploratory factor analysis, and shortened the questionnaire based on the psychometric properties of the emerging factors. We describe this process in this section.

4.1 Questionnaire Deployment Methods

Questionnaire validation requires substantial deployment and is most valid when tested with a diverse set of participants. In order to achieve this, we partnered with the Minnesota State Fair and set up a booth where participants could try out MST technologies and respond to the experience. 114 adults (57 pairs) participated in this study. The average age of participants was 37 (SD = 16.7) and 53% were female. We purposefully recruited pairs that knew each other (29 pairs), as well as pairs of strangers (28 pairs), to get a diversity of relationships.

All participants began by signing an IRB-approved consent form. All participants were asked to perform one of two 10-minute collaborative tasks with one of two MST systems. The Minnesota State Fair provided us with an opportunity to validate this instrument with a demographically-diverse group. An unfortunate trade-off was that the task needed to be relatively short to be appropriate for the State Fair context. We considered a number of possible tasks deciding to focus on a collaborative design task as one that could be appropriate between both strangers or known pairs, one that required substantial collaboration, and one that could feasibly be accomplished in a 10-minute session. We acknowledge that a different set of tasks could influence the specific dimensions of responses on the SDMST questionnaire and encourage this as an area for future work. Specifically, we chose two design tasks that were functionally similar but may involve different levels of affect:

- Discussing a scenario of a fictional company and collaboratively designing and drawing out a logo for it.
- Discussing a struggle from childhood (e.g., moving, failing a test) and collaboratively designing and drawing a poster to help children facing the same struggle.

We also wanted to validate the questionnaire with different types of technologies, so participants used one of two MST systems:

- ShareTable [47], which provides video chat and a shared tabletop duplexed projector-camera surface. Previous studies have shown that this system is interpreted as mediated touch by participants [52].
- ShareTable + Haptic Bands, which combines the ShareTable and a set of actuated hand and shoulder bands. These hand and shoulder bands leverage memory shape alloys to communicate a squeezing sensation and heat. The bands were triggered when certain interaction gestures were detected. For example, if both participants held their hands up in a "handshake" gesture, they would each feel their

[2] https://itunes.apple.com/us/app/couple-relationship-app-for-two/id503663173?mt=8

hand-band constrict and heat up to simulate touch. All gesture detection was done by Wizard-of-Oz to ensure optimal accuracy.

After completing the 10-minute tasks, participants responded to a number of metrics including our initial 49-question set, the Networked Minds Measure of Social Presence, and the NASA-TLX. For his or her help, each participant received a university-branded drawstring backpack. While we provide the necessary background on the study design for questionnaire validation, our other work provides more detail about the deployment setup and systems used [51]. We refer to this deployment in multiple sections of the paper, but first focus on the exploratory factor analysis.

4.2 Varimax Exploratory Factor Analysis

We conducted an exploratory factor analysis of the 114 responses to the 49 questions in order to identify factors. Our goal was to identify psychometrically optimal questions and scales to include on the questionnaire as a way of shortening and refining the question set. We performed a varimax (orthogonal) rotation, resulting in 11 factors having eigenvalues above one. The scree plot also confirmed that these were significant to examine. These informed our question and scale selection process. In the first round, questions were eliminated if they loaded onto multiple factors and had a low factor load. Factors were eliminated as potential scales if they contained fewer than three questions after this process. This reduced the question set from 49 to 38 questions and reduced the number of factors from 11 to five. From here, we conducted multiple iterations of removing low factor loading questions to optimize the inter-item reliability (Cronbach's alpha) of each factor. In order to optimize the inter-item reliability, we evaluated different combinations of questions by eliminating questions that had loads less than 0.6 and loaded onto multiple factors. Each time a question was eliminated, the result was documented in a summary table to help the research team observe patterns. Because removal of certain questions shifted the factor composition, it was important to document the effect of a question on each factor. One entire factor with 9 questions was eliminated because the Chronbach's alpha was consistently low, peaking at 0.7. Through this process, we reduced the question set from 38 to 21 questions and from five factors to three.

Given that experts in our formative questionnaire (see 3.2) reported using multiple validated metrics in a single study, we knew that questionnaire length would be an important factor in its adoption and use. Because of this, our goal was to reduce the number of questions needed to provide valid results. We ran the exploratory factor analysis with each factor having five, four, and three questions, looking to maximize the inter-item reliability, while minimizing the number of questions and retaining only those with the highest factor loading. We were able to achieve excellent inter-item reliability with just three questions per factor (see section 5.1). Further reducing the number of questions per scale would lower the inter-item reliability below acceptable parameters. We present the final set of questions, scales, and factor loadings in Table 2.

4.3 Final Factors and Scales

Retaining the most significant, orthogonal factors identified from the exploratory factor analysis, we named each of these factors as a scale. Although questions may seem redundant, all questions in a given scale measure a single construct. The three scales are: Social Discomfort, Communicational Expressiveness and Need for Additional Consideration. Section 4.4 expands on the definition and significance of these terms in the context of MST technologies.

Based on these factors, we were also able to more specifically name the resulting questionnaire—Social Disfordance of Mediated Social Touch (SDMST) questionnaire. We coin the term "social disfordance" (as the opposite of a social affordance [10]) to describe a situation where an MST technology hinders a social interaction in a particular context. Each scale and questions are listed in Table 2.

Table 2. Final factors and items with Cronbach's Alphas and factor loading. All items were tested with a 7-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree."

Scales & Questions	Factor Loading
Social Discomfort, □ *=0.936*	
The way [this technology] transmitted touch made me embarrassed to use it with [this partner].	0.920
The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].	0.912
The way [this technology] transmitted touch made me worry about making [this partner] uncomfortable.	0.898
Communicational Expressiveness, □ *=0.894*	
The way [this technology] transmitted touch contributed to my understanding of what [this partner] was communicating.	0.940
The way [this technology] transmitted touch contributed to my understanding of [this partner's] intentions.	0.922
The way [this technology] transmitted touch, I could express my intentions accurately to [this partner].	0.840
Need for Additional Consideration, □ *=0.876*	
I had to carefully consider how I would use [this technology] to transmit touch to [this partner].	0.908
I had to carefully interpret [this partner's] meaning in how they communicated touch using [this technology].	0.870
The way [this technology] transmitted touch required me to take extra care in how I communicated with [this partner].	0.842

4.4 Terms and Concepts

The SDMST questionnaire is a 9-item, three-scale questionnaire that examines how a particular technology used in a particular user context may elicit a positive or negative social experience. In this section, we explain the significance of each scale, as it relates to Social Disfordance.

Social Disfordance is a concept that focuses on the social aspects of a technology in a particular context. More specifically, social disfordance describes a situation where an MST technology lacks the affordances to support social interaction in a particular context. The SDMST questionnaire uses three scales to measure

different aspects of social disfordance: Social Discomfort, Communicational Expressiveness, and Need for Additional Consideration.

Social Discomfort is the extent to which this MST technology feels inappropriate or uncomfortable for the particular context of the social interaction. It addresses the fact that an MST technology may feel inappropriate or uncomfortable for a specific situation, taking away from an experience. The questions ask participant to reflect on how the technology made them feel when communicating with their partner. For example, in previous studies, similar concepts were measured with questions regarding the "awkwardness" or "invasiveness" of an interaction (e.g., [12]), though no validated measures of this concept exist. A high degree of Social Discomfort contributes to a higher Social Disfordance score.

Communicational Expressiveness measures the extent to which the MST technology supports the user in expressing their thoughts and feelings as intended. It looks at how the touch element of the MST technology helps users achieve a common ground for the communication. Similar concepts have been considered in previous literature, for example "Emotional Expressiveness" [50] and "Perceived Message Understanding" [19],

but previous instruments do not address the unique considerations of social touch signals. A high degree of Communicational Expressiveness contributes to a lower Social Disfordance score.

Need for Additional Consideration is the extent of extra social consideration or effort required to communicate via this technology in this particular context. Messages sent via certain technologies may be harder to interpret or may provide room for misinterpretation. In certain contexts (e.g., work), the misinterpretation of a touch message can be particularly problematic. Similar constructs have been considered in the framing of "user burden" in previous instruments [43], but approached individual time and emotional burdens, rather than aspect specific to the labor of communication. Technologies that require additional effort or consideration to send a message as intended have a higher degree of Social Disfordance in a particular context.

5. PSYCHOMETRIC PROPERTIES

The reliability and validity of the SDMST questionnaire were evaluated using the measures of inter-item reliability, convergent validity with existing instruments, test-retest reliability, and concurrent validity in the pilot study.

5.1. Inter-Item Reliability

Inter-item reliability measures the consistency of results across items (questions) within a test. This shows how consistent the answers are for each question for the same factor. We gathered 114 responses to a pilot version of the SDMST (see section 4.1). Measures of inter-item reliability are very sensitive to the number of questions per scale. Despite this, we were able to achieve Cronbach's alpha results ranging between "Excellent" (on Social Discomfort and Communication Expressiveness) and "Good" (on Need for Additional Consideration) with only three questions per scale (labels are heuristics offered in [9]). Based on these metrics, the SDMST meets the requirement of inter-item reliability for use.

5.2 Convergent Validity with Existing Instruments

During the deployment described in section 4.1, participants were also asked to take the NASA Task Load Index (NASA-TLX) scale and the Networked Minds Measure of Social Presence (NMMSP) questionnaire. We hypothesized that elements of these questionnaires would weakly or moderately correlate with certain factors on the SDMST, establishing convergent validity, which shows that measures that should be related are in reality related. Convergent validity, with these two other validated measures, establishes the validity of the SDMST questionnaire. The results are described below and summarized in Table 3.

Table 3. Convergent validity with existing validated measures, with a priori hypothesized correlations highlighted in color.

SDMST Scale	NASA-TLX	Perceived Message Understanding (NMMSP scale)	Perceived Affective Understanding (NMMSP scale)	Networked Minds Measure of Social Presence (full)
Social Discomfort	n.s	r = -0.207 ($p<.05^{*}$)	r = -0.297 ($p<.001^{***}$)	r = -0.243 ($p<.001^{***}$)
Communicational Expressiveness	r=-0.198 ($p<0.05^{*}$)	r = 0.242 ($p<.001^{***}$)	r = 0.364 ($p<.001^{***}$)	r = 0.362 ($p<.001^{***}$)
Need for Additional Consideration	r = 0.267 ($p<.01^{**}$)	n.s.	r = -0.195 ($p<.05^{*}$)	n.s.

We hypothesized that the strongest correlation with the NASA-TLX would be the Need for Additional Consideration factor because additional social effort would introduce a cognitive workload for the participant. Indeed, this correlation (Pearson *r*) was statistically significant ($r = 0.267$, $p < .01$). While we did not hypothesize other correlations with NASA-LTX, "Communicational Expressiveness" did display a weaker correlation ($r = 0.198$, $p < .05$). Perhaps, when a system supported greater communication expressiveness, the task became less difficult for participants. Overall, the NASA-TLX displayed expected convergence with the SDMST.

We also anticipated that the Perceived Message Understanding and Perceived Affective Understanding scales of the Networked Minds Measure of Social Presence would moderately correlate with the Communicational Expressiveness scale of the SDMST, because greater perceived expressiveness could lead to greater perceived understanding. Indeed, these correlations were statistically significant for both the Perceived Message Understanding ($r = 0.242$, $p < .001$) and the Perceived Affective Understanding ($r = 0.364$, $p < .001$) scales. We also saw a statistically significant inverse relationship between the social discomfort scale and both of the above scales (respectively, $r = -0.207$, $p < .05$ and $r = -0.297$, $p<.001$). We did not directly hypothesize this relationship a priori, but in retrospect a perceived lack of message and affective understanding could lead to social discomfort in a particular situation.

Outside of specific subscales, we hypothesized that there would be a negative correlation between Social Discomfort and the full Networked Minds Measure of Social Presence, as compelling social presence is harder to achieve in uncomfortable social interactions. Indeed, these two values were related in the expected and statistically significant way ($r = -0.243$, $p < .001$). Finally, we hypothesized that the "Communicational Expressiveness" scale would positively correlate with Networked Minds Measure of Social Presence since social presence implies the ability to meaningfully communicate the intended thought or feeling. Indeed, we found the expected strong positive relationship at the statistically significant level ($r = 0.362$, $p < .001$). Overall, the SDMST scales achieved expected convergence with the NMMSP.

SDMST achieved the majority of the anticipated convergences with related metrics. It is also important to note that the magnitude of all of each of these correlations is between "minor" and "moderate" (not exceeding ±0.5, a rule of thumb suggested by Cohen [9]). This suggests that while the SDMST converges with some expected concepts, none of these earlier metrics (alone or combined) are a sufficient replacement for the questionnaire designed in this work.

5.3 Test-Retest Reliability

Another way of demonstrating the reliability of a questionnaire is taking two measures separated by a time period and comparing the correspondence between the two results. This is a problematic measure for metrics that refer to a particular episode or experience (like the SDMST). One option is asking participants to take part in the experience once, but reflect on it twice (immediately after and two weeks after). This approach can introduce significant recall bias. The alternative is trying to set up two similar experiences separated by two weeks and having participants fill out the questionnaire immediately following each session. The challenge is that no two experiences are inherently the same, reducing the expected correlation between the two measures. This, combined with the difficulty of recruiting participants several weeks apart, is why many questionnaires validated in the HCI community omit this metric (e.g., [43]). Despite these difficulties and constraints, we thought that this was an important aspect of the validation.

We recruited a total of 10 students on the University of Minnesota campus to participate in two 10-minute tasks (similar to the deployment described in 4.1) using the ShareTable system and respond to the SDMST. Although 10 is considered a small sample size, we were limited by the number of participants available over a two-week period. The two tasks were separated by 14 days. Each participant provided consent and received a drawstring backpack for their participation. The partner pairing was kept consistent between studies (all pairs were "acquaintances"). We expected a strong and statistically significant correlation between these two measures, but perhaps a weaker one than psychometric testing of immutable quantities (e.g., personality, intelligence). Indeed, the correlation (Pearson r) of the Social Disfordance of

Mediated Social Touch questionnaire given two weeks apart was 0.649, which is statistically significant ($p < 0.01$). This magnitude of correlation is considered "high" for behavioral science studies, but only falls into the "acceptable" level for test-retest reliability [9]. It may have been "acceptable" rather than "high" for a number of reasons. First, participants got to know each other during the first trial and while waiting for the subsequent trial. Their increased level of familiarity may have influenced their scores. Similarly, participants may have become more familiar with the ShareTable as a communication medium through the first trial, again affecting the expressiveness and comfort they experienced with the task. Overall, the test-retest reliability of the SDMST is acceptable, but researchers should be wary of assuming that participants will retain scores across multiple trials. Counterbalancing order and accounting for learning effects may be particularly important with this measure.

5.4 Concurrent Validity

The concurrent validity of a questionnaire refers to its ability to distinguish between groups that should theoretically have substantially different responses. For example, in the pilot investigation, we would hypothesize that strangers using Mediated Social Touch systems will report substantially higher degrees of Social Disfordance than communication partners who knew each other prior to the study, because touch is not always appropriate between strangers. To demonstrate concurrent validity, the questionnaire scores would have to show a statistically significant difference between scores from strangers and scores from known pairs. We compared these groups using a General Linear Model at a .05 significance level. We found that this difference was statistically significant ($p=0.004$, $\eta^2=0.081$), confirming the test's initial concurrent validity. However, this is but one such hypothesis confirmed and we strongly encourage future investigations to contribute to our understanding of this measure's concurrent validity by deploying this instrument in diverse experimental contexts.

6. DISCUSSION

This section provides information on how to use and score the questionnaire. We discuss its inherent limitations and appropriate use

6.1 SDMST Questionnaire Scoring and Analysis

The Social Disfordance of Mediated Social Touch questionnaire is scored on a seven-point Likert scale (1 to 7), with Strongly Agree assigned to 7 and Strongly Disagree assigned to 1. We chose the 7-point Likert scale because bipolar constructs are best measured on an odd-point scale, given that the addition of a midpoint increases reliability, has no effect on validity, and does not result in lower data quality [31]. Once all of the answers are obtained, reverse score "Communicational Effectiveness" since the questions are negatively worded. For example, if a participant answered 7, the reverse score is 1, if the participant answered 6, the reverse score is 2 and so on. Now the score for all three scales, "Social Discomfort", Communicational Expressiveness, and "Need for Additional Consideration" are added to the overall Social Disfordance score. The scores may range from 3 to 54. A higher score signals a greater social disfordance of a particular MST technology, in a particular context, (in other words, a higher social disfordance score signified that the MST technology was not appropriate for the particular context). Table 4 shows an example of a scored questionnaire.

With descriptive investigations, scores on each scale can be aggregated across participants and presented as descriptive statistics (average and standard deviation). With hypothesis-driven work, sets of data can be compared using standard statistical tests for within-subjects or between-subject designs. A t-test or paired t-test is sufficient when the data is normally distributed. In cases where the normality assumption does not hold, we recommend the Wilcoxon rank sum and the Wilcoxon signed-rank tests as alternatives.

Table 4. Scoring guide example (in this case, the MST technology introduced low social disfordance to the particular experience)

Questions *(1 = strongly disagree, 7 = strongly agree)*	**Sample Answers**
Social Discomfort (SD, added score = 4)	
The way [this technology] transmitted touch made me embarrassed to use it with [this partner].	1
The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].	2
The way [this technology] transmitted touch made me worry about making [this partner] uncomfortable.	1
Communicational Expressiveness (CE, added score after reversal = 6)	
The way [this technology] transmitted touch contributed to my understanding of what [this partner] was communicating.	7 (reverse scored as 1)
The way [this technology] transmitted touch contributed to my understanding of [this partner's] intentions.	5 (reverse scored as 3)
The way [this technology] transmitted touch, I could express my intentions accurately to [this partner].	6 (reverse scored as 2)
Need for Additional Consideration (NAC, added score = 6)	
I had to carefully consider how I would use [this technology] to transmit touch to [this partner].	3
I had to carefully interpret [this partner's] meaning in how they communicated touch using [this technology]	2
The way [this technology] transmitted touch required me to take extra care in how I communicated with [this partner].	1
Total Social Disfordance Score:	**16**

6.2 Guidelines for SDMST Questionnaire Use

The SDMST questionnaire can be administered to adults on paper or online. Our pilot deployment with 114 adults was completed through a web interface, with minimal instruction for participants. The main value of the SDMST is in allowing the investigator to make comparisons between different technologies, tasks, and user populations. This is not an exhaustive list of acceptable investigations, but potential scenarios for appropriate use may include:

- Meta-Analysis Regarding a Particular Relationship: If an investigator wants to conduct a meta-analysis of SDMST scores for different MST technologies (e.g., remote hugging, haptic gloves, etc.) for a particular relationship type (e.g., parents and children), then he or she can gather all papers that report SDMST scores with those populations. The investigator can use standard meta-analysis techniques to compare and draw conclusions about the types of technologies that may be of most value to supporting MST in parent-child relationships.
- Between-Subject Comparison of Two Tech Versions: If an investigator wants feedback on a new iteration of a particular technology (e.g. hand shaking in a business negotiation setting), then he or she can deploy the new iteration and have participants fill out the questionnaire after using it. Then the investigator can compare the data with the previous iteration's data to understand the differences in the social disfordance of the two versions of the technology.
- Within-Subject Comparison of Two Tech Versions: If an investigator wants to test versions of a mediated social touch technology (like the ShareTable with and without haptic bands), he or she can

ask the same group of participants to try out and respond to the questionnaire about each system. The investigator can use pairwise comparison to understand whether the addition of the haptic bands increased or decreased the social disfordance for that particular relationship and task.

Flex-N-Feel [42] is one of the most recent MST prototypes in CSCW. This within-subject study compared Skype and Skype with a haptic glove using a variety of methods for data collection such as video-footage of the study, self-developed post-interview questionnaire, and an interview. We suggest that a 9-item questionnaire like the SDMST could be added to similar future investigations to allow for better comparison between these two technologies. If these results are reported, future investigators can combine similar findings across studies to better understand the role of haptics in augmenting videochat.

While there are a number of acceptable study designs involving the SDMST, we offer two significant cautions for its use. First of all, the SDMST questionnaire is not designed to serve as a sole metric for evaluating a novel MST technology. It captures some but not all of the important constructs of the user experience and **we strongly recommend that it be combined with other relevant measures and with rich qualitative methods** whenever possible. Second, we caution investigators against making between-study comparisons in a small dataset (if there are not enough papers on a particular topic for a meta-analysis). Particularly, **drawing favorable or unfavorable comparisons against a technology in one previous investigation is more likely to be misleading than informative**. The SDMST score is likely to be influenced not only by the specific MST system but also by the particular relationship type and by the type of task attempted by the participants. All of these factors should be clearly reported along with SDMST scores to support replication and meaningful future meta-analyses.

6.3. Limitations

When using the Social Disfordance of Mediated Social Touch questionnaire, researchers should understand its limitations. We developed this questionnaire to complement existing validated metrics already in use in the field. As such, the SDMST questionnaire should not be used in isolation as the only evaluation metric for a novel MST technology. It also does not replace the nuanced and rich data that can be provided with qualitative methods. The SDMST supports comparison across alternatives and the potential for meta-analysis across studies. However, triangulating data from rich qualitative methods and the structured data from validated questionnaires would lead to more valid and insightful results than using either method alone.

6.4. Adaptations of the SDMST Questionnaire Use

This paper is an initial step towards validated, replicable ways of evaluating Mediated Social Touch technologies; however, there can be a number of other efforts that would help the research community move towards this goal.

6.4.1. Experience Sampling. This questionnaire was developed for use in the lab, since that is the current setting of most MST studies. It can also be deployed at the completion of a field study of an MST technology, as the domain matures and field deployments become more common. However, given the momentary nature of social interaction, it may also be helpful to create adaptations of this instrument that could be used for continuous experience sampling in the field. While it is notoriously difficult to design notifications [11] or interruptions [36] for effective pervasive applications, mobile on-the-go evaluation [48] may provide the most reliable results in field deployments of MST systems. The short length of this questionnaire makes it amenable for splitting up into individual questions asked one at a time, for example, on the locked screen of a device [45]. Previously, adaptations of validated metrics delivered as single question at a time over several days have been shown to have similar reliability and validity as intact on-paper delivery of the same assessment (e.g., [53]). As the SDMST is adopted for evaluating Mediated Social Touch technologies, evaluating the use of a single question at a time over several days may be a reasonable next step to investigating its validity and reliability for more context-dependent, in-the-moment field evaluations.

6.4.2. Short-Form Version. Proper validation of questionnaires requires establishing inter-item reliability and therefore asking multiple questions to measure a single scale construct. Asking all three questions on the scale increases confidence that the reader correctly understood the underlying construct that is being investigated (in other words, it decreases individual variance in scale scores, which is a desirable characteristic). However, we acknowledge that in many HCI contexts, an even shorter set of questions may be strongly preferred (e.g., if the questionnaire is distributed after frequent short interactions or in a long multi-condition study). One approach that may be helpful to investigators in such a situation is taking just the single question with the highest factor loading from each scale. When taking this approach, authors should refer to this new instrument as the "adapted" or "short-form" version of the SDMST.

6.4.3. Adaptations for Other Modalities. Specificity is important in the initial validation of an instrument, so we focused exclusively on the touch modality in the SDMST. The current initial validation holds only for the touch modality. However, future work looks at modifying the questionnaire for other senses with further validation. We think such adaptations may require substantial revisions as the constructs in the questionnaire emerged from a systematic iterative review of MST work. However, the idea of "social disfordance" may be applied to other mediating technologies and the SDMST may serve as a starting point for such instrument design. If investigators modify the SDMST for other modalities, they should not claim that they are using the SDMST but rather say that they generated a novel instrument inspired by the SDMST.

7. CONCLUSIONS

This paper discusses the process of designing and validating a questionnaire to measure the social disfordance of Mediated Social Touch technologies. Three scales define the SDMST measure: Social Discomfort, Communicational Effectiveness, and Need for Additional Consideration. We present the development of the questionnaire through feedback from experts and deployment at the Minnesota State Fair with 114 participants. We optimized the questionnaire by using exploratory factor analysis for factor reduction. We demonstrate the reliability of the resulting instrument with inter-item reliability and test-retest reliability measures. We demonstrate its validity by reporting concurrent validity and expected and significant convergent validity with the NASA-TLX and the Networked Minds Measure of Social Presence questionnaire. Although the questionnaire has its limitations, it is a good step towards standardizing the way mediated social touch technologies are evaluated in research settings.

ACKNOWLEDGMENTS

We gratefully acknowledge the contributions on the GroupLens Research Center and its members in continually providing feedback and help with this design and validation process. We thank the Driven to Discover Research Facility at the Minnesota State Fair for providing us with the opportunity to carry out this study. Finally, we thank our funders: NSF (#1526085) and UMN SOBACO Faculty Award.

REFERENCES

[1] Leonardo Bonanni, Cati Vaucelle, Jeff Lieberman, and Orit Zuckerman. 2006. TapTap: A Haptic Wearable for Asynchronous Distributed Touch Therapy. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '06), 580–585. DOI:https://doi.org/10.1145/1125451.1125573

[2] S. Brave, H. Ishii, and A. Dahley. 1998. Tangible interfaces for remote collaboration and communication. In *Proc. of CSCW*, 169–178.

[3] Scott Brave and Andrew Dahley. 1997. inTouch: a medium for haptic interpersonal communication. In *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, 363–364. Retrieved February 15, 2017 from http://dl.acm.org/citation.cfm?id=1120435

[4] Scott Brave, Clifford Nass, and Erenee Sirinian. 2001. Force-Feedback in computer-mediated communication. In *HCI*, 145–149. Retrieved February 15, 2017 from https://www.researchgate.net/profile/Clifford_Nass2/publication/221099192_Force-Feedback_in_computer-mediated_communication/links/02bfe50cb681925e1e000000.pdf

[5] Stephen Brewster and Lorna M. Brown. 2004. Tactons: structured tactile messages for non-visual information display. In *Proceedings of the fifth conference on Australasian user interface-Volume 28*, 15–23. Retrieved February 15, 2017 from http://dl.acm.org/citation.cfm?id=976313

[6] Jongeun Cha, Mohamad Eid, Ahmad Barghout, A. S. M. Rahman, and Abdulmotaleb El Saddik. 2009. HugMe: synchronous haptic teleconferencing. In *Proceedings of the 17th ACM international conference on Multimedia*, 1135–1136. Retrieved February 15, 2017 from http://dl.acm.org/citation.cfm?id=1631535

[7] Angela Chang, Sile O'Modhrain, Rob Jacob, Eric Gunther, and Hiroshi Ishii. 2002. ComTouch: design of a vibrotactile communication device. In (DIS '02), 312–320. DOI:https://doi.org/10.1145/778712.778755

[8] Drew Cingel and Anne Marie Piper. 2017. How Parents Engage Children in Tablet-Based Reading Experiences: An Exploration of Haptic Feedback. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), 505–510. DOI:https://doi.org/10.1145/2998181.2998240

[9] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2 edition ed.). Routledge, Hillsdale, N.J.

[10] Alan Costall. 1995. Socializing Affordances. *Theory Psychol.* 5, 4 (November 1995), 467–481. DOI:https://doi.org/10.1177/0959354395054001

[11] Mary Czerwinski, Ran Gilad-Bachrach, Shamsi Iqbal, and Gloria Mark. 2016. Challenges for Designing Notifications for Affective Computing Systems. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (UbiComp '16), 1554–1559. DOI:https://doi.org/10.1145/2968219.2968548

[12] Andre Doucette, Regan L. Mandryk, Carl Gutwin, Miguel Nacenta, and Andriy Pavlovych. 2013. The Effects of Tactile Feedback and Movement Alteration on Interaction and Awareness with Digital Embodiments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13), 1891–1900. DOI:https://doi.org/10.1145/2470654.2466250

[13] M. A. Eid and H. Al Osman. 2016. Affective Haptics: Current Research and Future Directions. *IEEE Access* 4, (2016), 26–40. DOI:https://doi.org/10.1109/ACCESS.2015.2497316

[14] Mario Enriquez, Karon MacLean, and Christian Chita. 2006. Haptic phonemes: basic building blocks of haptic communication. In *Proceedings of the 8th international conference on Multimodal interfaces*, 302–309. Retrieved February 15, 2017 from http://dl.acm.org/citation.cfm?id=1181053

[15] Van Erp, Jan B. F, and Alexander Toet. 2015. Social Touch in Human–Computer Interaction. *Front. Digit. Humanit.* 2, (2015). DOI:https://doi.org/10.3389/fdigh.2015.00002

[16] Tiffany Field. 2003. *Touch. London: A Bradford Book*. The MIT Press, Massachusetts.

[17] Alberto Gallace and Charles Spence. 2010. The science of interpersonal touch: An overview. *Neurosci. Biobehav. Rev.* 34, 2 (February 2010), 246–259. DOI:https://doi.org/10.1016/j.neubiorev.2008.10.004

[18] Antal Haans and Wijnand IJsselsteijn. 2006. Mediated social touch: a review of current research and future directions. *Virtual Real.* 9, 2–3 (2006), 149–159.

[19] Chad Harms and Frank Biocca. 2004. Internal consistency and reliability of the networked minds measure of social presence. (2004). Retrieved February 15, 2017 from http://cogprints.org/7026/

[20] Sandra G. Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, 904–908.

[21] Marc Hassenzahl, Dr Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In *Mensch & Computer 2003*, Gerd Szwillus and Jürgen Ziegler (eds.). Vieweg+Teubner Verlag, 187–196. DOI:https://doi.org/10.1007/978-3-322-80058-9_19

[22] Marc Hassenzahl, Stephanie Heidecker, Kai Eckoldt, Sarah Diefenbach, and Uwe Hillmann. 2012. All You Need is Love: Current Strategies of Mediating Intimate Relationships Through Technology. *ACM Trans Comput-Hum Interact* 19, 4 (December 2012), 30:1–30:19. DOI:https://doi.org/10.1145/2395131.2395137

[23] Eve Hoggan, Roope Raisamo, and Stephen A. Brewster. 2009. Mapping information to audio and tactile icons. In *Proceedings of the 2009 international conference on Multimodal interfaces*, 327–334.

[24] Wijnand A. IJsselsteijn, Yvonne A. W. de Kort, and Antal Haans. 2006. Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence Teleoperators Virtual Environ.* 15, 4 (2006), 455–464.

[25] Stanley E. Jones and A. Elaine Yarbrough. 1985. A Naturalistic Study of the Meanings of Touch. *Commun. Monogr.* 52, 1 (1985), 19–56.

[26] M. M. Jung, R. W. M. Boensma, G. Huisman, and B. van Dijk. 2013. Touched by the Storyteller: The Influence of Remote Touch in the Context of Storytelling. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*, 792–797. DOI:https://doi.org/10.1109/ACII.2013.147

[27] Sara Kiesler, Jane Siegel, and Timothy W. Mcguire. 1984. Social psychological aspects of computer-mediated communication. *Am. Psychol.* 39, 10 (1984), 1123–1134. DOI:https://doi.org/10.1037/0003-066X.39.10.1123

[28] Jane Lessiter, Jonathan Freeman, Edmund Keogh, and Jules Davidoff. 2001. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence Teleoperators Virtual Environ.* 10, 3 (2001), 282–297. DOI:https://doi.org/10.1162/105474601300343612

[29] Margaret McLaughlin, Younbo Jung, Wei Peng, SeungA Jin, and Weirong Zhu. 2008. Touch in computer-mediated communication. *Mediat. Interpers. Commun.* (2008), 158.

[30] Ir van Mensvoort and Mfa Koert. 2002. What you see is what you feel: exploiting the dominance of the visual over the haptic domain to simulate force-feedback with cursor displacements. In *Proceedings of the 4th conference on Designing interactive systems: processes, practices, methods, and techniques*, 345–348.

[31] Hendrik Muller, Aaron Sedley, and Elizabeth Ferrall-Nunge. 2014. Survey Research in HCI. In *Ways of Knowing in HCI*, Judith S. Olson and Wendy A. Kellogg (eds.). Springer New York, New York, NY, 229–266. DOI:https://doi.org/10.1007/978-1-4939-0378-8_10

[32] Hideyuki Nakanishi, Kazuaki Tanaka, and Yuya Wada. 2014. Remote handshaking: touch enhances video-mediated social telepresence. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, 2143–2152.

[33] Ian Oakley, Stephen Brewster, and Philip Gray. 2001. Can you feel the force? An investigation of haptic collaboration in shared editors. In *Proceedings of EuroHaptics*, 54–59.

[34] H. Ogawa, N. Ando, and S Onodera. 2005. SmallConnection: designing of tangible communication media over networks. In *Proceedings of Mobile Media*, 1073–1074.

[35] Young-Woo Park, Kyoung-Min Baek, and Tek-Jin Nam. 2013. The roles of touch during phone conversations: long-distance couples' use of POKE in their homes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1679–1688.

[36] Veljko Pejovic and Mirco Musolesi. 2014. InterruptMe: Designing Intelligent Prompting Mechanisms for Pervasive Applications. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '14), 897–908. DOI:https://doi.org/10.1145/2632048.2632062

[37] Gregory R. Pierce. 1994. The Quality of Relationships Inventory: Assessing the interpersonal context of social support. (1994).

[38] A. F. Rovers and H. A. van Essen. 2006. Guidelines for haptic interpersonal communication applications: an exploration of foot interaction styles. *Virtual Real.* 9, 2–3 (March 2006), 177–191. DOI:https://doi.org/10.1007/s10055-005-0016-0

[39] James A. Russell. 1980. A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, (1980), 1161–1178.

[40] Bahador Saket, Chrisnawan Prasojo, Yongfeng Huang, and Shengdong Zhao. 2013. Designing an effective vibration-based notification interface for mobile phones. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 149–1504.

[41] Johnny Saldaña. 2015. *The coding manual for qualitative researchers*. Sage.

[42] Samarth Singhal, Carman Neustaedter, Yee Loong Ooi, Alissa N. Antle, and Brendan Matkin. 2017. Flex-N-Feel: The Design and Evaluation of Emotive Gloves for Couples to Support Touch Over Distance. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW '17), 98–110. DOI:https://doi.org/10.1145/2998181.2998247

[43] Hyewon Suh, Nina Shahriaree, Eric B. Hekler, and Julie A. Kientz. 2016. Developing and Validating the User Burden Scale: A Tool for Assessing User Burden in Computing Systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 3988–3999.

[44] Edmund R. Thompson. 2007. Development and validation of an internationally reliable short-form of the Positive and Negative Affest Schedule (PANAS).(Author abstract). *J. Cross-Cult. Psychol.* 38, 2 (2007), 227.

[45] Khai N. Truong, Thariq Shihipar, and Daniel J. Wigdor. 2014. Slide to X: Unlocking the Potential of Smartphone Unlocking. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems* (CHI '14), 3635–3644. DOI:https://doi.org/10.1145/2556288.2557044

[46] Dzmitry Tsetserukou. 2010. HaptiHug: A novel haptic display for communication of hug over a distance. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, 340–347.

[47] Baris Unver, Sarah A. McRoberts, Sabirat Rubya, Haiwei Ma, Zuoyi Zhang, and Svetlana Yarosh. 2016. ShareTable Application for HP Sprout. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 3784–3787.

[48] Heli Väätäjä and Virpi Roto. 2010. Mobile Questionnaires for User Experience Evaluation. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '10), 3361–3366. DOI:https://doi.org/10.1145/1753846.1753985

[49] Rongrong Wang, Francis Quek, Deborah Tatar, Keng Soon Teh, and Adrian Cheok. 2012. Keep in touch: channel, expectation and experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 139–148.

[50] Svetlana Yarosh, Panos Markopoulos, and Gregory D. Abowd. 2014. Towards a Questionnaire for Measuring Affective Benefits and Costs of Communication Technologies. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing* (CSCW '14), 84–96. DOI:https://doi.org/10.1145/2531602.2531634

[51] Svetlana Yarosh, Kenya Mejia, Baris Unver, Xizi Wang, Yuan Yao, Akin Campbell, and Bradley T. Holschuh. 2018. SqueezeBands: Mediated Social Touch Using Shape Memory Alloy Actuation. *PACM on Human-Computer Interaction* 1, 1 (2017), To Appear.

[52] Svetlana Yarosh, Anthony Tang, Sanika Mokashi, and Gregory D. Abowd. 2013. Almost touching: parent-child remote communication using the sharetable system. In *Proceedings of the 2013 conference on Computer supported cooperative work*, 181–192.

[53] Tae-Jung Yun, Hee Young Jeong, Tanisha D. Hill, Burt Lesnick, Randall Brown, Gregory D. Abowd, and Rosa I. Arriaga. 2012. Using SMS to Provide Continuous Assessment and Improve Health Outcomes for Children with Asthma. In *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium* (IHI '12), 621–630. DOI:https://doi.org/10.1145/2110363.2110432

Received April 2017; revised July 2017; accepted November 2017.