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

A THESIS SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Kenya Zarahi Mejia

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN MECHANICAL ENGINEERING

Svetlana Yarosh

June 2017

Acknowledgements

I'd like to acknowledge all of the people who contributed to the work that allowed the study at the Driven to Discover booth to become a reality. This includes both volunteers and contributors to the technologies used. I'd like to thank Sarah McRoberts, Baris Unver, Xizi Wang, Yuan Yao, Akin Cambell, Jackie Zamow, Julia Duvall, Brad Holschuh, Aaron Free, and Megan Clarke. We would also like to thank the experts who provided feedback on our initial categories.

Separately, I'd like to thank my advisor, Lana Yarosh, who inspired me through her passionate work. I'd also like to thank my defense committee. William Durfee, Barry Kudrowitz, and Julianna Abel, for their time and for their individual commitment to my growth and development as a scholar.

Dedication

This thesis is dedicated to my family and friends who believed in me and supported me through this process. To my parents, Maura and Luis who have worked hard so my sisters and I can focus on making our dreams a reality. To my sisters, Karen and Carol, who were always a phone call away. To Ian, my nephew, who is a reminder of the reasons I wake up and do what I do. And last but not least, to God who is my rock.

Abstract

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. This work discusses the design and validation 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 that resulted in an initial set of 49 questions to be deployed in a large-scale study. Its refinement included an exploratory factor analysis with a subsequent reduction of questions and scales. The final questionnaire, with three scales and three questions each , was created using the data from 114 participants. Included is a report of its psychometric properties, including metrics of inter-item reliability, convergent validity, and test-retest validity, confirming that these properties are sufficient for future use. It concludes with examples of scoring, appropriate use, a discussion of the limitations and future work. As is a limitation for many questionnaires, the SDMSTQ should be used in addition to other validated measures in order to help create a full evaluation of a system.

Table of Contents

List of Tables

List of Figures

Nomenclature

α	Chronbach's Alpha
MST	Mediated Social Touch
NASA TLX	NASA Task Load Index
NMMSP	Networked Minds Measure of Social Presence
r	Pearson's r
SDMST	Social Disfordance of Mediated Social Touch

1. Introduction

Touch is a fundamental part of human communication, showing its importance from the time one is born. 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 [1]. Touch is used to communicate emotions such as friendliness, affection, intimacy, support, playfulness, and more [2], but it can also be used to influence other's social behavior, such as modulating tendencies to comply with requests or affecting people's attitudes towards specific services [3]. Touch also has different meanings based on the given situation. In a situation where some stranger touches one, touch can illicit fear or discomfort, but a loved one's touch can bring a sense of contentment and security. As technology continues to develop, researchers are still looking at ways to transmit the sense of touch across long-distance communication. Currently, people can communicate across distance using the auditory senses (phones) and both their auditory and visual senses (video chat), but in order to fully capture face-to-face interactions, it will be important to appropriately replicate the sense of touch in technologies being developed. Because human perceptions of social touch reflect thousands of years of evolutionary development, and specific types of touch are suitable or unsuitable for specific relationships, contexts, and tasks [2], mediated social touch becomes a challenging design space.

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 people to touch across distance in order to reinforce social relationships [4]. With MST technologies, people can have the experience of sensing, transmitting, and simulating social touch between remote partners. Mediated social touch has been explored in research settings, but even so, little technology using MST has made it to the consumer market. As interest in communication technologies for enhancing social relationships grows, so does the development of new MST systems in research settings. 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., [5–9]). 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 [10]. 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 experience. In this work, a questionnaire instrument was developed and validated for this purpose.

Presented here is the development of a 9-item questionnaire that helps measure the "Social Disfordance" of mediated social touch technologies. Social disfordance is coined as the opposite of social affordance [11], where high social disfordance of a mediated social touch technology implies that the technology is a detriment to the sensing, transmitting and/or simulating of social touch in that particular context. 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. The three scales in the questionnaire are: Social Discomfort, Communicational Expressiveness, and Need for Additional Consideration. Finally, various metrics for validation confirmed that the SDMST questionnaire's psychometric properties are sufficient for use. Touch can have different meanings depending on the situation, therefore it is important for researchers to be able to evaluate their technologies and compare them with other technologies, previous iterations of the same technology, and even evaluate the same technology in different touch scenarios.

I begin by discussing related work associated with mediated social touch and questionnaire validation in Human Computer Interaction (HCI) and related fields. Described next is the process of designing and developing a questionnaire, starting with a systematic literature review and then reporting on how expert and pilot participant feedback was used to iterate and refine the questionnaire. Also discussed is the use of an exploratory factor analysis to reduce and optimize the number of questions. Following the exploratory, I report the questionnaire's validation methods, which includes inter-item reliability, convergent validity, and test-retest reliability tests. Finally, the discussion section provides guidance on questionnaire scoring, appropriate use, and reflections on the limitations of the SDMST and future work.

2. Related work

Haptic communication technologies and Mediated Social Touch (MST) have been of significant interest to the research community (e.g., [9,12–15]). This section discusses the major threads of research in both of these approaches, highlighting the difference between haptic communication and MST. Also discussed is the previous work done to measure the effects of MST technologies.

2.1. Haptic Technologies for Communication and Mediated Social Touch

The literature review revealed 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., [16]). Although this is targeting the sense of touch, the touch here is not being used to communicate with another person. 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 device (e.g., [6,17]). Other studies focused on developing and validating alternative alphabets for information transmission by developing units of meaning known as tactons [18], tactile icons [19], or haptic phonemes [20]. Another significant class of devices focuses on

"affective haptics" which attempt to represent and transmit specific emotions as tactile sensations (reviewed in [21]). Many of these affective haptic technologies are piloted with couples, limiting the scope of relationship type touch information that is gathered.

In contrast to haptic communication, 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. 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., [22–24]).

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 [25], 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. [26] 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. [14] 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 [27] built a vibrotactile vest and Cha et al [28] 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" [10] and largely "suffer from a lack of robust empirical testing" [3]. While these statements were made several years ago, MST work as recent at 2017 still does not rely on employing validated metrics for evaluation (e.g., [9]). 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. Although it is helpful and important for researchers to continue developing technologies that improve upon the current state of transmitting the sense of touch remotely, it will be important to be able to empirically evaluate these technologies to compare with previous

work and in different social scenarios, to improve these technologies in order to them to be ready for use by the masses. As Brave et al [29] acknowledges, the underlying sentiment is that adding a dimension of touch will lead to more efficient and satisfying interfaces, but through the study, shows that touch can have both a positive or negative effect. Again, it is important to be able to empirically test these theories through standardized metrics.

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 [30]. There are also a number of metrics for assessing the quality of any specific relationship, such as the Quality of Relationships Inventory [31]. However, these metrics are not 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). Although these metrics can be helpful in specific contexts, they are a not appropriate for the lab research setting in which many of these technologies are developed.

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., [32,33]) and in Human Computer Interaction (e.g., [22,34]). 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. It was 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., [35,36]). Researchers also develop their own questions to specifically address their technology and use interviews, sometimes using both pre and post interviews, as supplemental information. Unfortunately, this means that each study asks different questions and these questions are not validated in any way. Not only does this approach put the validity and reliability of these investigations in question, it also prevents combining results through systematic meta-analysis. Although the use of qualitative data is encouraged, having standardized, validated questionnaires will be of more help when comparing to other technologies or improved versions of the same technology. Additionally, it can help with time and effectiveness as the result is a statistical dataset. This investigation, presents 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

This section describes 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.

3.1. Systematic Literature Review

The completed systematic literature review, which included 103 relevant papers on mediated social touch started with the collecting of papers using the Google Scholar database. It was important to use incognito mode to prevent previous searches from influencing the results. The search included the phrases "tangible communication," "haptic communication," and "mediated social touch," and the results were sorted by relevance. The top 100 papers for each search term was evaluated. A research team read the abstracts of all resulting papers in order to develop an exclusion criteria and from the beginning, excluded papers that did not include technological mediation of touch and papers that focused on touch between robots and people (rather than social communication between two people). After applying the exclusion criteria to the initial set of 300 papers and removing duplicates, over one-third remained, leaving 103 total papers to be considered in the next round of evaluation. The papers were read and analyzed to be sorted according to the technologies, relationships, and situations that each

study used to address mediated social touch. More relevant to this investigation though, through reading the papers and pulling out questions used in each study, a list of all of these questions asked in the various studies were compiled. In some cases, the questions were not reported in the paper and the original authors were contacted in order to get the complete list of questions asked (there was success in connecting with all the authors contacted).

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

Question Category	Previous Work Example Question
Intensity	How overwhelmed, or calm, did you feel when being greeted in this manner?
Physical Realism	How realistic, or unrealistic, did you find this greeting? (Very unrealistic to Very realistic)
Value Added	How well were you able to express this emotion [with this MST system]?
Social Comfort	What did you perceive as comfortable or uncomfortable during the experiment, if anything?

There was quite a bit of diversity in wording and topics 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, all of the questions were clustered based on similarity of topic or underlying concept. Three researchers collaborated in clustering the questions

into general themes and refining the clusters through three separate passes until a clear set of categories emerged.

Through this clustering, several clusters of questions were identified 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. It was found that a number of categories of questions attempted to measure aspects of systems that already have existing validated metrics. For example, one category of questions related to social presence. These types of questions were excluded from subsequent analysis, as there already exist validated metrics for measuring social presence (e.g., Networked Minds Measure of Social Presence- NMMSP [22]). The categories of questions excluded included ones 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 Task Load Index or NASA-TLX [37].

In the remaining dataset, questions fell into four categories. (Table 1 provides examples of the questions from each of these four major categories):

- **Intensity:** was this interaction intense or calm?

- **Physical Realism:** was this interaction similar to touching in an unmediated way?
- **Value Added:** did the touch aspects of the system enhance the communication experience?
- **Social Comfort:** was the way this system transmitted touch natural and socially appropriate for the relationship and context?

These categories represented the questions that were being asked in research settings to evaluate MST systems, but did not have corresponding standardized wording or measures that could be used in future analysis or comparison. In order to provide a questionnaire that would be relevant to the people who would be using it, it was crucial to contact experts in the field for feedback on these categories.

3.2. Expert Feedback

For the purpose of this study, any author of a peer-reviewed paper on MST was considered to be an expert on the topic. The systematic literature review was used to identify potential experts, which resulted in the identification and contacting of the 148 authors of the 103 papers included in the 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 identified above, Intensity, Physical Realism, Value Added, and Social Comfort. The experts were asked to rate each category's importance to evaluating MST technologies (on a five-point Likert-type scale of importance) and comment as to whether any important aspects of MST systems were missing from the current analysis. The experts were also provided with 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). These questions can be found in Appendix A. Fourteen of the authors representing 13.4% of the contacted authors and 18.4% papers (several of them were authors on multiple papers) responded and provided feedback.

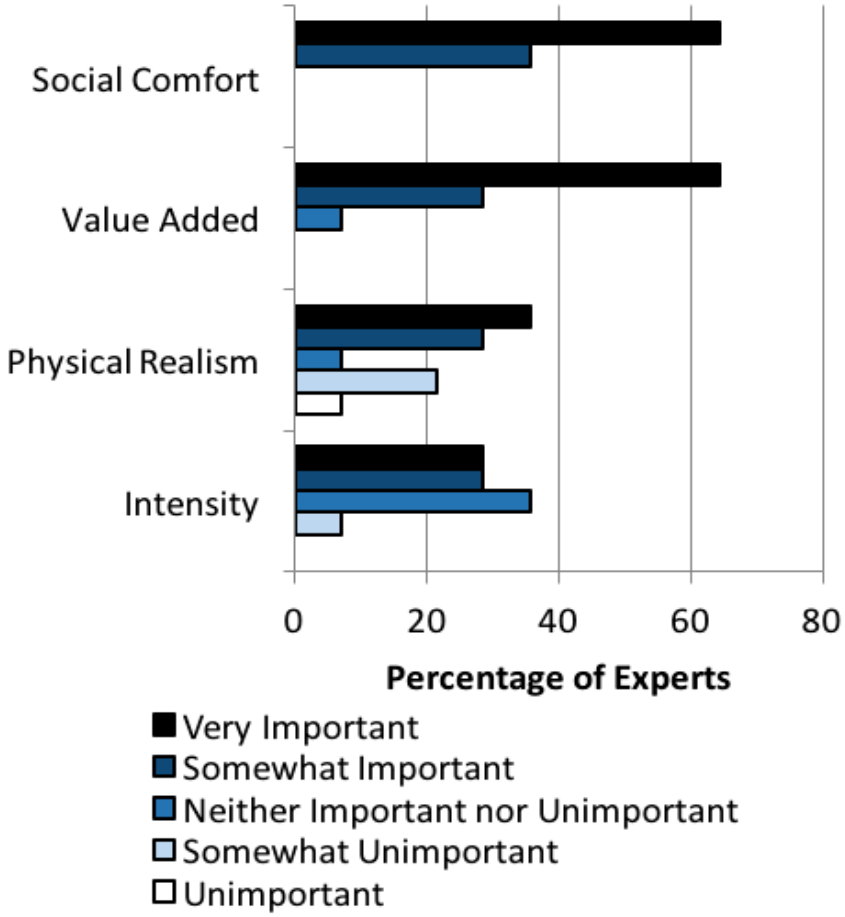

Figure 1. Shows the responses from the experts. The graph shows the categories where experts are in agreement and where there is a range of opinions.

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 it to be pursued in this 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.

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 [38]) 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, for the rest of the study commensurate focus was given to these ideas when designing the questionnaire. The other two categories were not pursued.

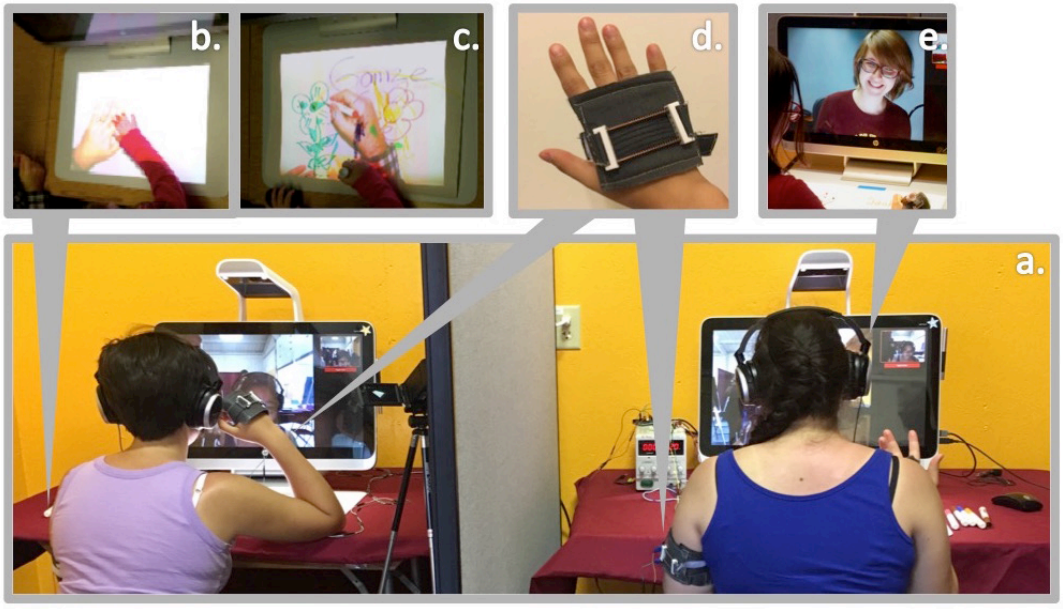

Figure 2. The study setup during the deployment: (a) paired participants sat at tables separated by a wall panel; (b) duplexed projector-camera surface allows for mediated social touch through overlapping video, as well as (c) provides a shared work space; (d) haptic hand and shoulder bands respond to touch gestures with squeezing and heat; (e) the video screen shows a standard face-to-face video chat view and feedback window.

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. At the end of a brainstorming session, 52 questions that were inspired by expert feedback made up the final set for piloting. These questions 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.

The following step was to make sure these questions would make sense when asked in the final deployment of the questionnaire. 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."[1] 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. Instead of just answering the brainstormed questions from above, they were asked to also 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 later to aid in the analysis of these transcripts to identify which questions were most problematic. This phase concluded 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, three questions were eliminated, all of which all three participants found confusing, and modified six questions, which some of the participants found difficult to understand, to address points of potential confusion in the future of the study. The refined set of 49 questions, documented in Appendix B, was used in the next phase.

4. Exploratory Factor Analysis and Questionnaire Refinement

In order to continue refining and validating the questionnaire, the questionnaire was deployed to 114 adult participants, an exploratory factor analysis was conducted, and the questionnaire was shortened based on the psychometric properties of the emerging

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

factors. The process is described in this section.

4.1. Questionnaire Deployment Methods

Questionnaire validation requires substantial deployment and is most valid when tested with a diversity of participants. In order to achieve this, the ProDUCT lab, through the University of Minnesota, partnered with the Minnesota State Fair and set up a booth where participants could try out MST technologies and respond to the experience using existing validated questionnaires and the refined set of 49 SDMST questions. 114 adults (57 pairs) participated in this study. The average age of participants was 37 (SD = 16.7) and 53% were female. Participants were purposefully recruited as 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 and a short survey to gather demographic information. All participants were asked to perform one of two 10-minute collaborative tasks with one of two MST systems. Table 2 shows how many pairs were recruited for each of the eight potential scenarios. The tasks were intentionally chosen to be functionally similar but each one potentially involved different levels of affect:

- Scenario 1: Discussing a scenario of a fictional company and collaboratively designing and drawing out a logo for it.

- Scenario 2: 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.

The goal was also to validate the questionnaire with different types of technologies in, so participants used one of two MST systems (see Figure 2):

- System 1: ShareTable [39], 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 [15].

- System 2: 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 help their hands up in a "handshake" gesture, they would each feel their hand-band constrict and heat up to simulate touch. All gesture detection was done by Wizard-of-Oz to ensure optimal accuracy.

Table 2. Breakdown of the number of pairs that participated in each of the eight possible scenarios

		Stranger Pair	Known Pair
System 1	Low Emotion	9	8
	High Emotion	6	7
System 2	Low Emotion	7	7
	High Emotion	6	7

After completing the 10-minute tasks, participants responded to a number of metrics including the 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 the discussion of these specific systems and their comparative results are not discussed here, they provided useful examples for validating the initial questionnaire. The rest of the sections refer to this deployment in multiple sections, but the next section will focus on the exploratory factor analysis done after the deployment.

4.2 Varimax Exploratory Factor Analysis

An exploratory factor analysis of the 114 responses to the 49 questions was conducted in order to identify factors. Using exploratory factor analysis is useful in identifying relationships or factors among variables, or in this case, questions. The goal was to identify psychometrically optimal questions and scales to include on the questionnaire as a way of shortening and refining the question set. It was important to keep the questionnaire as short as possible to make it a quick tool useful in a research setting

where there is usually limited time with participants and many tasks to be performed or questions to be answered. A varimax (orthogonal) rotation was performed, resulting in 11 factors having eigenvalues above one. The scree plot, shown in figure 3, also confirmed that these factors were significant to examine.

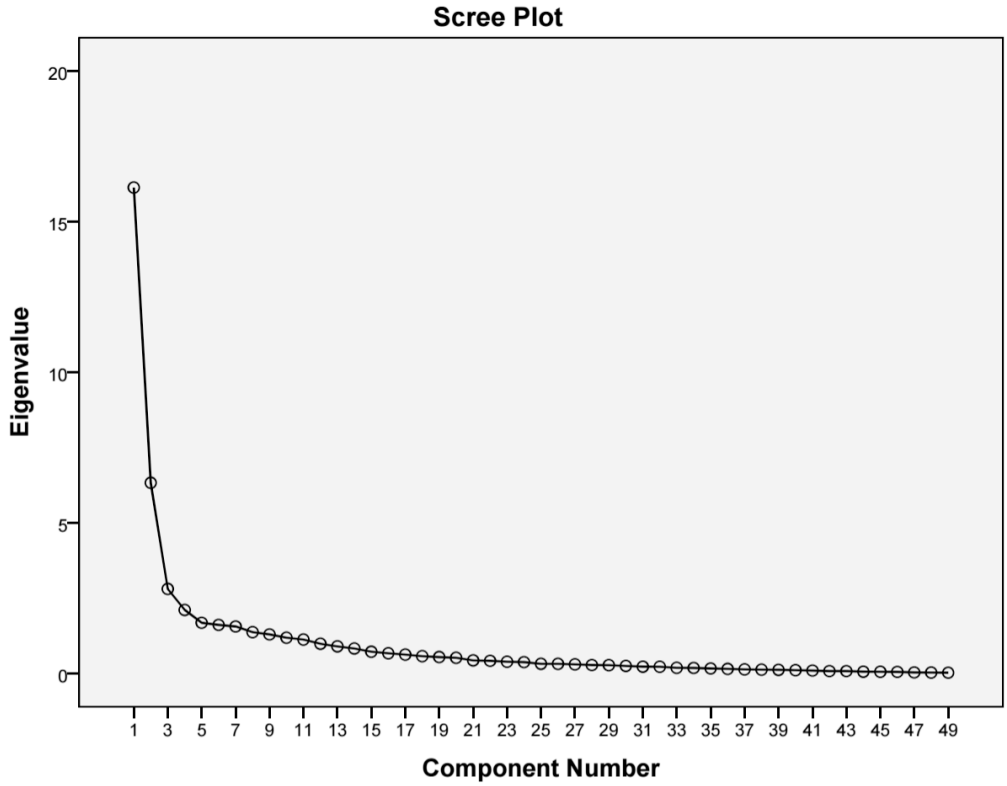

Figure 3. The scree plot shows it is acceptable to use exploratory factor analysis to aid in the reduction of factors.

The scree plot and eigenvalues informed the question and scale selection process. In the first round, questions were eliminated if they loaded onto multiple factors and had a low factor load, implying that the question was not a strong fit in any of the factors it was

loading onto. Factors were eliminated as potential scales if they contained fewer than three questions after this process because it was not a substantial set of questions to define this factor. This reduced the question set from 49 to 38 questions and reduced the number of factors from 11 to five.

From here, multiple iterations were conducted 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, different combinations of questions were evaluated 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 researcher observe patterns. Table 3 is an example of how questions were eliminated or brought in to the exploratory factor analysis to define factor two. The data for the rest of the factors can be found in Appendix C. Because removal of certain questions shifted the factor composition, it was important to document the effect of a question on each factor.

Factor two, composed of 9 questions, was eliminated because the Chronbach's alpha was consistently low, peaking at 0.707. Through this process, the question set was reduced from 38 to 21 questions and from five factors to three. Given that experts in the formative questionnaire (see 3.2) reported using multiple validated metrics in a single study, it was known that questionnaire length would be an important factor in its adoption and use.

Table 3. Demonstrates the method used to reduce the number of factors from five to three by methodically eliminating questions to optimize the Chronbach's alpha, α, value of each factor, henceforth increasing the inter item reliability. The black boxes symbolize questions that were removed from that specific run and light grey boxes represent questions that had not yet been incorporated into that factor, but later became relevant as other questions were eliminated.

Factor	Questions	1st Run	Cronbach's Alpha (α)	2nd Run	Cronbach's Alpha (α)	3rd Run	Cronbach's Alpha (α)
Factor 2	The way [this technology] transmitted touch felt natural to use with [this partner].	0.744	0.072	0.819	0.707	0.794	0.524
	[This technology] felt like a natural way to communicate touch to [this partner].	0.72		0.82		0.799	
	Communicating touch to [this partner] using [this technology] felt like second nature to me.	0.719		0.815		0.811	
	It didn't take any extra effort to use [this technology] to communicate touch to [this partner].	0.707		0.73		0.763	
	It took a lot of extra effort to use [this technology] to communicate touch to [this partner].	-0.681		-0.627		-0.677	
	The way [this technology] transmitted touch made me worry about offending [this partner].	0.521					
	The way [this technology] transmitted touch was different from how I would convey touch to [this partner] in-person.	-0.492					
	The way [this technology] transmitted touch to [this partner] made it natural to communicate touch in this situation.			0.597			
	The way [this technology] transmitted touch was consistent with how I would convey touch to [this partner] in-person.			0.462			

Because of this, the goal was to reduce the number of questions needed to provide valid results. The exploratory factor analysis was performed once more, 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. An excellent inter-item reliability score was achieved with just three questions per factor (see section 5.1). The final set of questions, scales, and factor loadings are available in Table 4.

4.3. Final Factors and Scales

Retaining the most significant, orthogonal factors identified from the exploratory factor analysis, each of these factors were named as a scale. The three scales are defined as follows:

- **Social Discomfort:** the extent to which this MST technology feels inappropriate or uncomfortable for the particular context of the social interaction.

- **Communicational Expressiveness:** the extent to which the MST technology supports the user in expressing their thoughts and feelings as intended.

- **Need for Additional Consideration:** the extent of extra social consideration or effort required to communicate via this technology in this particular context.

Table 4. 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, *α=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, *α=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, *α=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

Based on these factors, it was also possible to more specifically name the resulting questionnaire—Social Disfordance of Mediated Social Touch Questionnaire (SDMST). Here the term "social disfordance" is coined (as the opposite of a social affordance [11]) to describe a situation where an MST technology hinders a social interaction in a

particular context. Social disfordance of a particular MST technology is increased if it is inappropriate or uncomfortable for a particular context (Social Discomfort scale); Social Disfordance is increased if the partners must take additional care or effort to communicate or interpret (Need for Additional Consideration scale); but, Social Disfordance is decreased if it supports the partners in accurately expressing their thoughts and feelings (Communicational Expressiveness scale). Each scale and questions are listed in Table 4.

5. Psychometric properties

The SDMST was validated on measures of inter-item reliability, convergent validity with existing instruments, and test-retest reliability. Each of these tests support the validity of the questionnaire.

5.1. Inter-Item Reliability

From the study, 114 responses were gathered for 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, the Cronbach's alpha results achieved still ranged between "Excellent" (on Social Discomfort and Communication Expressiveness) and "Good" (on Need for Additional Consideration) scales with only three questions per scale (labels are heuristics offered in [40]). As mentioned in section 4.2, the inter-item reliability was crucial in the decision of reducing factors, as each time there was a change in the set of questions, the Chronbach's alpha was calculated to determine whether or not that set of questions would yield a valid questionnaire. 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. It was hypothesized that elements of these questionnaires would weakly or moderately correlate with certain factors on the SDMST, establishing convergent validity. The results are described below and summarized in Table 5.

Table 5. 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.

It was 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 it was not hypothesized that other

correlations with NASA-LTX would occur, “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.

It was 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. The results also showed a statistically significant inverse relationship between the social discomfort scale and both of the above scales (respectively, $r = -0.207$, $p < .05$ and $r = -0.297$, $p<.001$). It was not directly hypothesized a priori that this relationship would occur, but in retrospect a perceived lack of message and affective understanding could lead to social discomfort in a particular situation.

Outside of specific subscales, it was also 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, it was 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, the results showed the expected strong positive relationship at the statistically significant level ($r = 0.362$, $p < .001$). Overall, the SDMST scales also 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 [40]). 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 by 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 challenge, combined with the difficulty of recruiting participants several weeks apart, is why many questionnaires validated in the HCI community omit this metric (e.g., [41]). Despite these difficulties and constraints, this metric was an important aspect of the validation that had to be included. A total of 10 students on the University of Minnesota campus were recruited to participate in two 10-minute tasks (similar to the tasks completed in the deployment described in 4.1) using the ShareTable system and respond to the SDMST. 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").

The expectation was 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).

Table 6. 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 = 18)	
The way [this technology] transmitted touch contributed to my understanding of what [this partner] was communicating.	7
The way [this technology] transmitted touch contributed to my understanding of [this partner's] intentions.	5
The way [this technology] transmitted touch, I could express my intentions accurately to [this partner].	6
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
Social Disfordance Score (SD – CE + NAC)	**-8**

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 validity [40]. 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.

6. Discussion

This section provides information on how to use and score the questionnaire. Also discussed are its inherent limitations and appropriate use scenarios.

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. Once all of the scores are obtained, the scores for "Social Discomfort" and "Need for Additional Consideration" scale are added to the overall Social Disfordance

score and the scores for "Communicational Effectiveness" are subtracted from that score. The final questionnaire can be found in Appendix D. The scores may range from -15 to 39. The overall score gives the social disfordance of a technology. A more positive score signals a greater social disfordance of a particular MST technology in a particular context. Table 6 shows an example of a scored questionnaire. The example shown shows the total for a low disfordance score.

With descriptive investigations, scores on each scale can be aggregated across participants and presented as descriptive statistics (average and standard deviation/error). 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, the Wilcoxon rank sum and the Wilcoxon signed-rank tests are recommended alternatives.

6.2. Guidelines for SDMST Questionnaire Use

The SDMST questionnaire can be administered to adults on paper or online. A copy of the full questionnaire is available in Appendix D. The 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. 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.

- Meta-Analysis Regarding the Particular Emotionality of a Task: If an investigator wants to conduct a meta-analysis of SDMST scores in studies ranging in emotionality of tasks (i.e. professional handshake, hugging, teaching, etc.), then he or she can gather all papers that report SDMST scores specific types of tasks, either defined by the research paper or determined by the investigator. He or she can use standard meta-analysis techniques to compare and draw conclusions about the effects of emotionality of tasks on the performance of the MST technology on the SDMST.

- Meta-Analysis Regarding a Particular Population: An investigator wants to conduct a meta-analysis of SDMST scores for the same system but used by different populations (e.g., in different cultures, younger vs. older generations). He or she gathers all papers that include the population(s) of interest. Given that these papers presented SDMST for this population, the investigator can use standard meta-analysis techniques to draw comparisons and conclusions about the types of communication technologies that may be of most value to this population and find differences among each population compared in the meta-analysis.

This is not an exhaustive list of acceptable investigations. However, investigators should take caution 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. This questionnaire was developed 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 nuance 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 rather than using either method alone. The questionnaire developed completes a portion of the puzzle.

6.4. Future Directions for Field Evaluations

This work 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. 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 [42] or interruptions [43] for affective pervasive applications, mobile on-the-go evaluation [44] 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 lock 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., [46]). As the SDMST is adopted for evaluating Mediated Social Touch technologies, this may be a reasonable next step to investigating its validity and reliability for more context-dependent, in-the-moment field evaluations.

7. Conclusions

This work 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. The work presents the development of the questionnaire through feedback from experts and deployment at the Minnesota State Fair with 114

participants. Using the feedback from experts, the questionnaire meets the need for analyzing the social disfordance of a mediated social touch technology. Although the questionnaire has its limitations, it is a good step towards standardizing the way mediated social touch technologies are evaluated in research settings.

References

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

[2] Jones, S. E., and Yarbrough, A. E., 1985, "A Naturalistic Study of the Meanings of Touch.," Commun. Monogr., **52**(1), pp. 19–56.

[3] Gallace, A., and Spence, C., 2010, "The Science of Interpersonal Touch: An Overview," Neurosci. Biobehav. Rev., **34**(2), pp. 246–259.

[4] Brave, S., Nass, C., and Sirinian, E., 2001, "Force-Feedback in Computer-Mediated Communication.," *HCI*, pp. 145–149.

[5] Bonanni, L., Vaucelle, C., Lieberman, J., and Zuckerman, O., 2006, "TapTap: A Haptic Wearable for Asynchronous Distributed Touch Therapy," *CHI '06 Extended Abstracts on Human Factors in Computing Systems*, ACM, New York, NY, USA, pp. 580–585.

[6] Chang, A., O'Modhrain, S., Jacob, R., Gunther, E., and Ishii, H., 2002, "ComTouch: Design of a Vibrotactile Communication Device," ACM, pp. 312–320.

[7] Jung, M. M., Boensma, R. W. M., Huisman, G., and Dijk, B. van, 2013, "Touched by the Storyteller: The Influence of Remote Touch in the Context of Storytelling," *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)*, pp. 792–797.

[8] Ogawa, H., Ando, N., and Onodera, S., 2005, "SmallConnection: Designing of Tangible Communication Media over Networks.," *Proc. of Mobile Media*, pp. 1073–1074.

[9] Singhal, S., Neustaedter, C., Ooi, Y. L., Antle, A. N., and Matkin, B., 2017, "Flex-N-Feel: The Design and Evaluation of Emotive Gloves for Couples to Support Touch Over Distance," *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ACM, New York, NY, USA, pp. 98–110.

[10] Haans, A., and IJsselsteijn, W., 2006, "Mediated Social Touch: A Review of Current Research and Future Directions," Virtual Real., **9**(2–3), pp. 149–159.

[11] Costall, A., 1995, "Socializing Affordances," Theory Psychol., **5**(4), pp. 467–481.

[12] Brave, S., Ishii, H., and Dahley, A., 1998, "Tangible Interfaces for Remote Collaboration and Communication," *Proc. of CSCW*, pp. 169–178.

[13] Cingel, D., and Piper, A. M., 2017, "How Parents Engage Children in Tablet-Based Reading Experiences: An Exploration of Haptic Feedback," *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, ACM, New York, NY, USA, pp. 505–510.

[14] Nakanishi, H., Tanaka, K., and Wada, Y., 2014, "Remote Handshaking: Touch Enhances Video-Mediated Social Telepresence," *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, ACM, pp. 2143–2152.

[15] Yarosh, S., Tang, A., Mokashi, S., and Abowd, G. D., 2013, "Almost Touching: Parent-Child Remote Communication Using the Sharetable System," *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, ACM, pp. 181–192.

[16] Saket, B., Prasojo, C., Huang, Y., and Zhao, S., 2013, "Designing an Effective Vibration-Based Notification Interface for Mobile Phones," *Proceedings of the 2013 Conference on Computer Supported Cooperative Work*, ACM, pp. 149–1504.
[17] McLaughlin, M., Jung, Y., Peng, W., Jin, S., and Zhu, W., 2008, "Touch in Computer-Mediated Communication," Mediat. Interpers. Commun., p. 158.
[18] Brewster, S., and Brown, L. M., 2004, "Tactons: Structured Tactile Messages for Non-Visual Information Display," *Proceedings of the Fifth Conference on Australasian User Interface-Volume 28*, Australian Computer Society, Inc., pp. 15–23.
[19] Hoggan, E., Raisamo, R., and Brewster, S. A., 2009, "Mapping Information to Audio and Tactile Icons," *Proceedings of the 2009 International Conference on Multimodal Interfaces*, ACM, pp. 327–334.
[20] Enriquez, M., MacLean, K., and Chita, C., 2006, "Haptic Phonemes: Basic Building Blocks of Haptic Communication," *Proceedings of the 8th International Conference on Multimodal Interfaces*, ACM, pp. 302–309.
[21] Eid, M. A., and Osman, H. A., 2016, "Affective Haptics: Current Research and Future Directions," IEEE Access, **4**, pp. 26–40.
[22] Harms, C., and Biocca, F., 2004, "Internal Consistency and Reliability of the Networked Minds Measure of Social Presence."
[23] IJsselsteijn, W. A., de Kort, Y. A. W., and Haans, A., 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), pp. 455–464.
[24] van Mensvoort, I., and Koert, M., 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," *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques*, ACM, pp. 345–348.
[25] Brave, S., and Dahley, A., 1997, "inTouch: A Medium for Haptic Interpersonal Communication," *CHI'97 Extended Abstracts on Human Factors in Computing Systems*, ACM, pp. 363–364.
[26] Park, Y.-W., Baek, K.-M., and Nam, T.-J., 2013, "The Roles of Touch during Phone Conversations: Long-Distance Couples' Use of POKE in Their Homes," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 1679–1688.
[27] Tsetserukou, D., 2010, "HaptiHug: A Novel Haptic Display for Communication of Hug over a Distance," *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, Springer, pp. 340–347.
[28] Cha, J., Eid, M., Barghout, A., Rahman, A. S. M., and El Saddik, A., 2009, "HugMe: Synchronous Haptic Teleconferencing," *Proceedings of the 17th ACM International Conference on Multimedia*, ACM, pp. 1135–1136.
[29] Brave, S., Nass, C., and Sirinian, E., 2001, "Force-Feedback in Computer-Mediated Communication.," *HCI*, pp. 145–149.

[30] Thompson, E. R., 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), p. 227.
[31] Pierce, G. R., 1994, "The Quality of Relationships Inventory: Assessing the Interpersonal Context of Social Support."
[32] Kiesler, S., Siegel, J., and Mcguire, T. W., 1984, "Social Psychological Aspects of Computer-Mediated Communication," Am. Psychol., **39**(10), pp. 1123–1134.
[33] Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J., 2001, "A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory," Presence Teleoperators Virtual Environ., **10**(3), pp. 282–297.
[34] Yarosh, S., Markopoulos, P., and Abowd, G. D., 2014, "Towards a Questionnaire for Measuring Affective Benefits and Costs of Communication Technologies," *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM, New York, NY, USA, pp. 84–96.
[35] Oakley, I., Brewster, S., and Gray, P., 2001, "Can You Feel the Force? An Investigation of Haptic Collaboration in Shared Editors," *Proceedings of EuroHaptics*, pp. 54–59.
[36] Wang, R., Quek, F., Tatar, D., Teh, K. S., and Cheok, A., 2012, "Keep in Touch: Channel, Expectation and Experience," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp. 139–148.
[37] Hart, S. G., 2006, "NASA-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Sage Publications Sage CA: Los Angeles, CA, pp. 904–908.
[38] Hassenzahl, M., Burmester, D. M., and Koller, F., 2003, "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität," *Mensch & Computer 2003*, G. Szwillus, and J. Ziegler, eds., Vieweg+Teubner Verlag, pp. 187–196.
[39] Unver, B., McRoberts, S. A., Rubya, S., Ma, H., Zhang, Z., and Yarosh, S., 2016, "ShareTable Application for HP Sprout," *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ACM, pp. 3784–3787.
[40] Cohen, J., 1988, *Statistical Power Analysis for the Behavioral Sciences*, Routledge, Hillsdale, N.J.
[41] Suh, H., Shahriaree, N., Hekler, E. B., and Kientz, J. A., 2016, "Developing and Validating the User Burden Scale: A Tool for Assessing User Burden in Computing Systems," *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM, pp. 3988–3999.
[42] Czerwinski, M., Gilad-Bachrach, R., Iqbal, S., and Mark, G., 2016, "Challenges for Designing Notifications for Affective Computing Systems," *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, ACM, New York, NY, USA, pp. 1554–1559.
[43] Pejovic, V., and Musolesi, M., 2014, "InterruptMe: Designing Intelligent Prompting Mechanisms for Pervasive Applications," *Proceedings of the 2014 ACM*

International Joint Conference on Pervasive and Ubiquitous Computing, ACM, New York, NY, USA, pp. 897–908.

[44] Väätäjä, H., and Roto, V., 2010, "Mobile Questionnaires for User Experience Evaluation," *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, ACM, New York, NY, USA, pp. 3361–3366.

[45] Truong, K. N., Shihipar, T., and Wigdor, D. J., 2014, "Slide to X: Unlocking the Potential of Smartphone Unlocking," *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, ACM, New York, NY, USA, pp. 3635–3644.

[46] Yun, T.-J., Jeong, H. Y., Hill, T. D., Lesnick, B., Brown, R., Abowd, G. D., and Arriaga, R. I., 2012, "Using SMS to Provide Continuous Assessment and Improve Health Outcomes for Children with Asthma," *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium*, ACM, New York, NY, USA, pp. 621–630.

Appendix A: Sample Questions Given to Experts for feedback

Intensity

Evaluates affective arousal induced by the experience of using the technology.
Participants will respond on a five point Likert scale.
[Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree]

Question

- [ ] My experience with [this technology] was intense.
- [ ] My experience with [this technology] was relaxed. (negatively scored)
- [ ] My experience with [this technology] was calm. (negatively scored)
- [ ] My experience with [this technology] was vigorous.
- [ ] My experience with [this technology] was harsh.

Physical Realism

Evaluates how close the technology simulates real life communication and interactions.
Participants will respond on a five point Likert scale.
[Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree]

Question

- [ ] My experience with [this technology] was realistic.
- [ ] My experience with [this technology] felt unnatural. (negatively scored)
- [ ] My experience with [this technology] was true to real-life.
- [ ] My experience with [this technology] felt artificial. (negatively scored)
- [ ] My experience with [this technology] was familiar.

Value Added

Evaluates whether the technology adds anything to the communication experience.
Participants will respond on a five point Likert scale.
[Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree]

Question

- [ ] [This technology] helped my communication experience.
- [ ] My experience with [this technology] was valuable.
- [ ] [This technology] was important to my communication experience.
- [ ] My communication experience was better with [this technology].
- [ ] My experience with [this technology] was satisfactory.

Social Comfort

Evaluates the social affective aspects of the technology.
Participants will respond on five point Likert scale.
[Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree]

Question

- [ ] My experience with [this technology] was awkward. (negatively scored)
- [ ] My experience with [this technology] was comfortable.
- [ ] My experience with [this technology] was weird. (negatively scored)
- [ ] My experience with [this technology] was appropriate.
- [ ] My experience with [this technology] was eerie. (negatively scored)

Appendix B: Original Set of Questions

Original set of Questions	
1	The way [this technology] transmitted touch was appropriate for my relationship with [this partner].
2	The way [this technology] transmitted touch to [this partner] was appropriate for this relationship
3	The way [this technology] transmitted touch was inappropriate for my relationship with [this partner].
4	The way [this technology] transmitted touch to [this partner] was inappropriate for this relationship.
5	The way [this technology] transmitted touch made me worry about offending [this partner].
6	The way [this technology] transmitted touch made me worry about making [this partner] uncomfortable.
7	The way [this technology] transmitted touch made me embarrassed to use it with [this partner].
8	The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].
9	The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].
10	I felt comfortable with the way [this technology] transmitted touch to [this partner].
11	I felt uncomfortable with the way [this technology] transmitted touch to [this partner].
12	I felt embarrassed to use [this technology] to transmit touch to [this partner].
13	I felt awkward when [this partner] communicated touch using [this technology].
14	I felt uncomfortable when [this partner] communicated touch using [this technology].
15	I felt embarrassed when [this partner] communicated touch using [this technology].

16	My communication with [this partner] was supported by the way [this technology] transmits touch.
17	My communication with [this partner] was disrupted by the way [this technology] transmits touch
18	The way [this technology] transmitted touch obstructed my communication with [this partner].
19	The way [this technology] transmitted touch distracted from my communication with [this partner].
20	The way [this technology] transmitted touch contributed to my understanding of [this partner's] intentions.
21	The way [this technology] transmitted touch detracted from my understanding of [this partner's] intentions.
22	The way [this technology] transmitted touch contributed to my understanding of what [this partner] was communicating.
23	The way [this technology] transmitted touch detracted from my understanding of what [this partner] was communicating.
24	The way [this technology] transmitted touch interfered with my communication with [this partner].
25	The way [this technology] transmitted touch supported my communication with [this partner].
26	I felt awkward using [this technology] to communicate touch to [this partner].
27	I could interpret the touch [this partner] transmitted using [this technology].
28	I could understand the touch [this partner] transmitted using [this technology].
29	The way [this technology] transmitted touch made it easy to understand what [this partner] was trying to convey.
30	The way [this technology] transmitted touch made me worry that [this partner] would misinterpret my intention.
31	The way [this technology] transmitted touch made me worry that I would send the wrong message to [this partner].
32	The way [this technology] transmitted touch, I could express my intentions accurately to [this partner].

33	The way [this technology] transmitted touch made it clear what I meant to express to [this partner]
34	The way [this technology] transmitted touch required me to take extra care in how I communicated with [this partner].
35	I had to carefully consider how I would use [this technology] to transmit touch to [this partner].
36	I had to carefully interpret [this partner's] meaning in how they communicated touch using [this technology]
37	I intentionally avoided [this partner's] touch as it was transmitted by [this technology].
38	I intentionally sought out [this partner's] touch as it was transmitted by [this technology].
39	It didn't take any extra effort to use [this technology] to communicate touch to [this partner].
40	It took a lot of extra effort to use [this technology] to communicate touch to [this partner].
41	[This technology] felt like a natural way to communicate touch to [this partner].
42	Communicating touch to [this partner] using [this technology] felt like second nature to me.
43	The way [this technology] transmitted touch felt natural to use with [this partner].
44	The way [this technology] transmitted touch made me worry about offending [this partner].
45	[This partner] and I had to agree on how to interpret the way [this technology] transmitted touch
46	[This partner] and I had to agree on how to communicate touch given the way it was transmitted by [this technology].
47	The way [this technology] transmitted touch was different from how I would convey touch to [this partner] in-person.
48	The way [this technology] transmitted touch was consistent with how I would convey touch to [this partner] in-person.
49	The way [this technology] transmitted touch to [this partner] made it natural to communicate touch in this situation.

Appendix C: Process or eliminating questions and factors

Factor	Questions	1st Run	Cronbach's Alpha	2nd Run	Cronbach's Alpha	3rd Run	Cronbach's Alpha
Factor 1	The way [this technology] transmitted touch made me embarrassed to use it with [this partner].	0.879	0.957	0.866	0.944	0.871	0.943
	The way [this technology] transmitted touch made me worry about making [this partner] uncomfortable.	0.844		0.859		0.861	
	The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].	0.842		0.824		0.839	
	I felt embarrassed to use [this technology] to transmit touch to [this partner].	0.821		0.824		0.813	
	The way [this technology] transmitted touch made me worry about offending [this partner].	0.81		0.821		0.813	
	The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].	0.774		0.756		0.763	
	I felt embarrassed when [this partner] communicated touch using [this technology].	0.768		0.799		0.797	
	I felt awkward when [this partner] communicated touch using [this technology].	0.703		0.751		0.739	
	I felt uncomfortable when [this partner] communicated touch using [this technology].	0.699		0.734		0.742	
	The way [this technology] transmitted touch to [this partner] was inappropriate for this relationship.	0.675					
	The way [this technology] transmitted touch made me worry that I would send the wrong message to [this partner].	0.667					
	The way [this technology] transmitted touch made me worry that [this partner] would misinterpret my intention.	0.623					
	I felt uncomfortable with the way [this technology] transmitted touch to [this partner].	0.533		0.464		0.457	
	The way [this technology] transmitted touch was inappropriate for my relationship with [this partner].	0.512		0.548			

	I felt awkward using [this technology] to communicate touch to [this partner].	0.504		0.734			
Factor 2	The way [this technology] transmitted touch felt natural to use with [this partner].	0.744	0.072	0.819	0.707	0.794	0.524
	[This technology] felt like a natural way to communicate touch to [this partner].	0.72		0.82		0.799	
	Communicating touch to [this partner] using [this technology] felt like second nature to me.	0.719		0.815		0.811	
	It didn't take any extra effort to use [this technology] to communicate touch to [this partner].	0.707		0.73		0.763	
	It took a lot of extra effort to use [this technology] to communicate touch to [this partner].	-0.681		-0.627		-0.677	
	The way [this technology] transmitted touch made me worry about offending [this partner].	-0.521					
	The way [this technology] transmitted touch was different from how I would convey touch to [this partner] in-person.	-0.492					
	The way [this technology] transmitted touch to [this partner] made it natural to communicate touch in this situation.			0.597			
	The way [this technology] transmitted touch was consistent with how I would convey touch to [this partner] in-person.			0.462			
Factor 3	The way [this technology] transmitted touch contributed to my understanding of what [this partner] was communicating.	0.807	0.922	0.852	0.902	0.813	0.897
	The way [this technology] transmitted touch contributed to my understanding of [this partner's] intentions.	0.768		0.81		0.789	
	The way [this technology] transmitted touch supported my communication with [this partner].	0.726		0.792		0.744	
	My communication with [this partner] was supported by the way [this technology] transmits touch.	0.658		0.628		0.617	

	The way [this technology] transmitted touch made it clear what I meant to express to [this partner]	0.543					
	The way [this technology] transmitted touch, I could express my intentions accurately to [this partner].	0.537					
	I could interpret the touch [this partner] transmitted using [this technology].					0.811	
	I could understand the touch [this partner] transmitted using [this technology].					0.735	
Factor 4	I had to carefully consider how I would use [this technology] to transmit touch to [this partner].	0.822	0.876	0.823	0.876	0.837	0.876
	I had to carefully interpret [this partner's] meaning in how they communicated touch using [this technology]	0.807		0.81		0.827	
	The way [this technology] transmitted touch required me to take extra care in how I communicated with [this partner].	0.799		0.826		0.818	
Factor 5	The way [this technology] transmitted touch to [this partner] was appropriate for this relationship	0.833	0.848	0.776	0.85	0.835	0.848
	The way [this technology] transmitted touch was appropriate for my relationship with [this partner].	0.826		0.841		0.858	
	I felt comfortable with the way [this technology] transmitted touch to [this partner].	0.504		0.468		0.561	

Appendix D: Questionnaire for use

	Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
Social Discomfort							
1. The way [this technology] transmitted touch made me embarrassed to use it with [this partner].	o	o	o	o	o	o	o
2. The way [this technology] transmitted touch made me uncomfortable to use it with [this partner].	o	o	o	o	o	o	o
3. The way [this technology] transmitted touch made me worry about making [this partner] uncomfortable.	o	o	o	o	o	o	o
Communicational Expressiveness:							
4. The way [this technology] transmitted touch contributed to my understanding of what [this partner] was communicating.	o	o	o	o	o	o	o
5. The way [this technology] transmitted touch contributed to my understanding of [this partner's] intentions.	o	o	o	o	o	o	o
6. The way [this technology] transmitted touch, I could express my intentions accurately to [this partner].	o	o	o	o	o	o	o
Need for Additional Consideration							
7. I had to carefully consider how I would use [this technology] to transmit touch to [this partner].	o	o	o	o	o	o	o
8. I had to carefully interpret [this partner's] meaning in how they communicated touch using [this technology]	o	o	o	o	o	o	o
9. The way [this technology] transmitted touch required me to take extra care in how I communicated with [this partner].	o	o	o	o	o	o	o

Each question is score from 1-7. Once all of the scores are obtained, add the scores for "Social Discomfort" and "Need for Additional Consideration" to the overall Social Disfordance score and subtract the scores for "Communicational Effectiveness" that score. The scores may range from -15 to 39.