

Towards a Questionnaire for Measuring Affective Benefits and Costs of Communication Technologies

Svetlana Yarosh
AT&T Research Labs
Florham Park, NJ, USA
lana@research.att.com

Panos Markopoulos
Eindhoven University of Technology
Eindhoven, The Netherlands
p.markopoulos@tue.nl

Gregory D. Abowd
Georgia Institute of Technology
Atlanta, GA, USA
abowd@cc.gatech.edu

ABSTRACT

As CSCW creates and investigates technologies for social communication, it is important to understand the emotional benefits and costs of these systems. We propose the Affective Benefits and Costs of Communication Technologies (ABCCT) questionnaire to supplement traditional qualitative methods of understanding communication media. We discuss the pilots of this survey with 45 children and 110 adults measuring the inter-item reliability of this instrument. We present the results of interviews with 14 children and 14 adults, which help confirm that the ABCCT measures the same constructs that may emerge through interview investigations. Finally, we demonstrate that the ABCCT is sensitive enough to discriminate between different communication technologies and has shown promise in some of its early adoption. Though the ABCCT is not without limitations, it may provide a way to compare technologies in field deployments, draw findings across investigations, and quantify the impact of specific design decisions.

Author Keywords

Questionnaire; CMC; communication; affect; evaluation.

ACM Classification Keywords

H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces – *Evaluation/Methodology*.

1. INTRODUCTION

Social communication frequently relies on generic communication media originally intended for work- and task-related purposes. Communication for most families relies on the telephone, text messaging, and videoconferencing [21]. At CSCW, interest has been growing in technologies that are aimed specifically at supporting affect-oriented communication (e.g., [27]). Relevant innovations may involve off-the-shelf technologies like tablets (e.g., [35]) or browsers or special purpose devices that enable informal social interaction (e.g., [43]). As the field progresses from early design explorations to robust technologies intended to address specific separation contexts, empirical evaluation of

the emerging user experience can benefit from opportunities to triangulate findings from both qualitative and quantitative measures. In this work, we present the initial design and validation of a quantitative instrument that may supplement the existing qualitative approaches.

Field deployments of interactive systems allow for “exploring user understandings, practices and the eventual uses” of the novel system [7]. Typically, researchers log systems use (e.g., messages sent through the system) and interview the users post deployment. This approach provides rich qualitative data about contextualized user experiences, usage patterns, and appropriation of the technology, prompting and guiding subsequent improvements of the system. However, relying exclusively on qualitative methods presents some problems for this domain. One of the main problems is that a lack of agreed-upon metrics makes it difficult to draw conclusions across investigations, deployments, and systems (as discussed in [27,51]). It is difficult to compare systematically a specific technology either to earlier technologies or to alternate versions with different subsets of features. We propose that some of these challenges can be addressed by adopting a common approach for measuring the costs and benefits of communication systems, in addition to the commonly used qualitative methods.

We present the Affective Benefits and Costs of Communication Technologies (ABCCT) questionnaire as one such potential approach to the problem of evaluating communication technologies, which provides the following benefits:

- It addresses a specific technology in question, rather than connectedness or relationship quality in general.
- It is applicable to both synchronous and asynchronous communication technologies.
- It has less than 30 items and takes less than 10 minutes to complete.
- It can be used with both children and adults, with preliminary reliability and validity metrics provided.
- It highlights both emotional benefits and emotional costs that a system may introduce.
- It is sensitive enough to show differences between two communication systems.

The goal of this paper is to describe the process of developing and testing the ABCCT, as well as its early adoption in research. We developed the ABCCT questionnaire by building extensively on previous work, though collaborations with experts in the field, and through an analysis of

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CSCW'14, February 15–19, 2014, Baltimore, Maryland, USA.
Copyright © 2014 ACM 978-1-4503-2540-0/14/02...\$15.00.
DOI string from ACM e-form confirmation email/see text of message.

interviews with separated parents and children (Section 4). In Sections 6 and 7, we describe our method of validating the survey and our findings, which show that the ABCCT is reliable and valid according to several metrics taken in piloting the questionnaire. In Section 8, we share early uptake of this instrument in the HCI and CSCW communities.

If you are an investigator considering the ABCCT for your own work, Section 5 contains a description of the scales of the ABCCT, Table 2 contains the full text of the questions, and Section 9 contains guidelines for analyzing the ABCCT, its appropriate use, and its limitations.

2. MOTIVATING AND RELATED WORK

We provide a review of current evaluation approaches for communication systems and briefly discuss previous efforts to develop validated metrics in this space.

2.1. Evaluations of Social Communication Technologies

We identified 35 papers published in archival HCI venues (e.g., CHI, CSCW) that presented and evaluated novel communication technologies for families, friends, or couples. As criteria for selecting papers, we looked at work published between 2000 and 2011 that presented a novel social communication technology and some form of evaluation. Here, we present the conclusions from this literature survey on the current practices in evaluating novel social communication technologies.

A significant proportion of papers in the HCI domain present novel communication technology ideas (e.g., [11,43]) or even novel systems (e.g., [45,47]) but do not provide any sort of evaluation. In this analysis, we included only those papers that provide a discussion of a system evaluation. Of the 35 papers we identified that provided an evaluation, 10 papers (29%) reported lab study evaluations. Most papers used passive observations and interviews in the evaluation, though 2 papers [34,54] also employed structured observation using specific validated criteria. Eleven papers (32%) involved field deployments of 1–4 weeks. All of these collected logs of system use and conducted periodic interviews with participants. Thirteen papers (38%) involved field deployments of 5–12 weeks. Most of these collected logs of system use and interviews with participants; two of these also collected diaries of participants' experiences. This demonstrates the prevalence of qualitative methods for evaluation in this field.

Only five of the studies we found included structured pre- and post-deployment questionnaires to understand the impact of system's use on the social relationships of the study participants. This is surprising, because investigations of family communication in the social science domain quite frequently involve the use of validated questionnaires in addition to other methods [28]. Kaye et al. stated that “participants were asked to subjectively rate the effectiveness and level of intimacy of their current methods of communication” [20]. No other detail was provided on the questions used or the effectiveness of these metrics. Rowan & Mynatt

[38] “attempted to measure changes of perceived awareness and connectedness by asking [participants] to rank these indicators using a Likert scale daily in their diary.” Unfortunately, this metric did not yield useful results as they saw a ceiling effect. Khan et al. took measures before the introduction of their experimental system to support intra-family awareness, after a week of use, and a week after the end of this intervention [22]. They did not apply any validated instrument for the evaluation. Romero et al. report a two week deployment of an experimental system in which they compared measures prior to and after a week of using their system [37]. For the quantitative evaluation they used the Affective Benefits and Costs questionnaire (ABC-Q), a predecessor to the instrument presented here, that was developed for the purposes of that study [19]. Subsequent applications of the ABC-Q have shown it to be difficult to understand and to apply in relation to diverse technologies, making its revision necessary. We present a reworking of the affective benefits and costs questionnaire, validating the resulting questionnaire with a more rigorous quantitative approach, and extending it for use with children.

We do not advocate that qualitative methods should be replaced or abandoned, but rather support a mixed methods approach. We agree with Brown et al. that field trials present inherent challenges that could be ameliorated through diversifying the methods and approaches used in understanding system use [7]. Qualitative methods provide rich descriptions of system use and point to potential areas of improvements. On the other hand, collecting quantitative and structured data about how the system is perceived through the use of questionnaires better supports comparison between pre- and post-deployment experiences, between multiple novel systems, and between different versions of the same technology. This may be particularly important in the field of communication technology, since participant numbers tend to be fairly small in these evaluations (with two recent notable exceptions [24,35]). In order to be able to draw meaningful conclusions across systems and deployments, and eventually allow generalization of conclusions, it is helpful to use standard methods that would allow combining and comparing data across investigations. Our hope is that the ABCCT can provide one such standardized approach.

2.2. Questionnaire Instruments for Evaluating Communication Technologies

There are a number of survey instruments in social science that one might consider using to evaluate affective costs and benefits of communication technologies. Noller & Feeney provide a review of validated questionnaires for measuring characteristics of family communication [28]. A person's general positive and negative affect can be measured using the Positive and Negative Affect Schedule [46]. Further, there are a number of metrics for assessing the quality of any specific relationship, such as the Quality of Relationships Inventory [33]. Such measures could be taken before and after the introduction of a novel technology to

evaluate its impact. However, these instruments measure characteristics that remain fairly stable over time. There is no evidence that any of these instruments are sensitive enough to detect a subtle change in a relationship, communication pattern, or personal affect that may be caused by a 1–12 week usage of a novel communication system, which is typical for field-tests in this domain.

There have been attempts within HCI literature to design survey instruments specifically aimed at evaluating communication technologies. Traditionally these focus on synchronous media, especially those intending to emulate and substitute physical co-presence. Validated inventories have been developed by Short, Williams, and Christie [42] and more recently the Networked Minds Social Presence questionnaire [15]. The limitations of such instruments pertain to the operationalization of social presence, which is not applicable for asynchronous systems. The Social Connectedness Questionnaire (SCQ) was developed to measure social connectedness as one particular benefit of communication technologies [3]. The SCQ can be used for comparing pre- and post- measures to reveal potential benefits of using a newly introduced communication technology. However, SCQ only evaluates one specific benefit of a new technology and is silent about potential downsides, which are crucial to understanding a design. Recognizing this need, IJsselsteijn et al. developed the Affective Benefits and Costs Questionnaire (ABC-Q) which provides the initial inspiration for the ABCCT [19]. However, the only validation of the ABC-Q was the measurements collected from 11 participants in a single case study deployment of novel technology [37]. Though its results were not in conflict with the qualitative data collected through diaries and interviews, no quantitative measure of agreement was provided. ABC-Q has been used in several research groups and industry projects; however, it is challenging for both respondents and designers. For example, the ABC-Q consists of 60 items, which makes it too long for most studies; and ABC-Q mixes questions regarding specific technologies with general contact questions, making it difficult to use it to compare specific technologies.

These early efforts at developing and validating surveys in HCI highlight that currently available instruments from the social science domain frequently do not meet the needs of HCI researchers. The testing of novel instruments in HCI is frequently limited to inter-item reliability statistics gathered from fairly small deployments. One of our goals in this paper is to provide an evaluation of a questionnaire that instead focuses on its convergent validity with qualitative metrics and its sensitivity to detecting changes between alternative communication media.

3. QUESTIONNAIRE DESIGN REQUIREMENTS

We wanted to design a questionnaire that could meet the needs of a designer in the field. Several requirements drove the design process of the ABCCT. The ABCCT must:

- Measure both the emotional benefits and emotional costs of communication technologies
- Be quick to administer, taking participants less than 15 minutes to answer
- Refer to a specific communication medium, since most of the time it will be used to compare technologies
- Demonstrate reliability and validity on multiple metrics
- Be sensitive enough to detect differences between communication technologies
- Be applicable to both synchronous and asynchronous communication technologies

For our own work in studying families, we had one additional requirement that the ABCCT must be validated for use with children as well as adults, since children are frequently partners in social communication but their voices are often excluded from the evaluation. In the next section, we describe our process in creating the ABCCT.

4. QUESTIONNAIRE DESIGN

We give an overview of the process used to generate and pilot the initial draft of the ABCCT.

Table 1. Original ABCCT Scales and Support for Each Scale in Previous Literature

Scale	Evidence for Scale Inclusion
Emotional Expressiveness	[2][12][17][23][30][36][53]
Engagement & Playfulness	[1][2][9][10][11][32][37][49][53]
Presence-in-Absence	[1][18][26][36][37][39][43][53]
Opportunity for Personal Effort	[17][23][36][37][41][44]
Opportunity for Social Support	[9][13][23][26][30][36][41][53]
Feeling Obligated	[17][37][39][44][47][53]
Unmet Expectations	[8][36][39][47][50]
Threat to Privacy	[5][8][26][37][44][47][53]

4.1. Generating the Scales

We created the children's ABCCT first and then adapted it for use by adults. Creating a survey appropriate for children forced us to simplify wording and make the questions more clear. Four types of previous work were used to draft ABCCT scales and questions: the ABC-Q [19], the theoretical background provided in related literature (see Table 1) re-coding of the interviews from a previously conducted study of parent-child communication in divorced families [53], and brainstorming by two social connectedness experts. This process yielded 12 scales with a total of 90 items. Review and discussion of the resulting measure by 2 experts lead to eliminating 23 questions and rewriting a significant portion of the remaining questions. This resulted in a draft of the survey with 8 scales and 55 items. Table 1 lists these scales and literature consulted at the time of the instrument's creation that supports the inclusion of each

scale. We used Borger & Hox's [6] recommendations to choose a Likert-type scale that focused on frequencies of events in a specific period, rather than agreement (e.g., "never" to "always," rather than "strongly disagree" to "strongly agree"). This guideline let the questionnaire to remain more concrete and avoided advanced vocabulary (e.g., "neutral" is considered a 5th grade-level reading word) therefore allowing the ABCCT to be more appropriate for use with children.

4.2. Piloting the Questionnaire

We followed Bell's guidelines in designing and testing the children's questionnaires, by conducting comprehensive pretesting, including cognitive interviewing [4]. To pilot the survey with children, 3 nine- and ten-year-old boys were asked to complete the ABCCT-Child one-on-one with the researcher while reasoning out loud. By observing when the children became restless while taking the survey, we noted that it needed to be cut in size by half. The next draft of the survey restructured two of the scales and eliminated 9 questions that the children found unclear. The resulting measure was piloted with another four children using a think-aloud protocol. From this pilot, the next draft had most questions refined and redundant ones eliminated, resulting in an 8-scale, 32 item questionnaire. Following, an additional deployment (see Section 6.1), factorial analysis was done on the survey results leading to one item being moved to a different scale and to the merging of two closely-related scales. Ten questions were omitted in an effort to shorten the survey and remove confusing questions. The resulting draft consisted of the 7-scale, 22-item ABCCT-Child questionnaire (Table 2).

A pilot was conducted with 8 adults, asking them to complete the draft ABCCT-Child one-on-one with the researcher while reasoning out loud. This led to a number of changes that resulted in the adult version of the questionnaire:

- Reciprocity items were added to appropriate scales, as suggested by [19].
- Questions were reworded and used more nuanced vocabulary, on the suggestion of the participants.
- One question was removed, as it was the only negatively phrased question, frequently leading adult participants to mark the opposite answer.

This resulted in a 7-scale, 26-item ABCCT-Adult (Table 2).

5. ABCCT SCALES AND QUESTIONS

Through this design process, the ABCCT was focused and refined to measure the 4 major affective benefits and 3 major affective costs of communication technologies.

Benefit 1: Emotional Expressiveness. One benefit of a communication technology may be the ability to share emotional states with a partner. This includes both being able to express one's own feelings and being able to perceive the feelings expressed by the partner.

Benefit 2: Engagement & Playfulness. Social communication is frequently characterized by playful exchanges and other attempts at making the communication more engaging for both partners. The ABCCT asks whether the communication via a certain medium is fun and exciting to the participant. Adults may also be sensitive as to whether the communication is engaging to their partner, so the adult version includes a reciprocity item.

Benefit 3: Presence-In-Absence. Presence-In-Absence is defined as the "subjective sense of social others whilst separated from them by time or space" [18]. Some communication technologies may be more suited to fostering feelings of "closeness" or "connectedness" to a partner, as well as better supporting specific behaviors like remembering a contact or frequently thinking about each other.

Benefit 4: Opportunity for Social Support. One of the benefits of communication technologies is the opportunity to provide social support without being physically present. This includes a general sense of the other person "being there" for you, reducing negative affect (such as soothing anxiety), and increasing positive affect (such as feeling "special" or loved). This may also include more instrumental forms of support such as giving advice or helping make a decision. The ABCCT-Adult includes items about both providing and about receiving social support, while the children's version focuses on receiving support.

Cost 1: Feeling Obligated. A communication technology may introduce an unwanted obligation to connect. This may include feeling compelled to communicate or respond when you would rather not do so and even feelings of guilt over not maintaining adequate communication. The adult version of the questionnaire also includes a reciprocity item, since adults are often sensitive of situations where they might be creating obligations for others.

Cost 2: Unmet Expectations. A communication technology may increase expectations for communication. When these expectations are not met, participants may experience negative affect. Expectations may include the frequency of contact, how quickly a person returns a missed contact, and how much attention is devoted to the communication. The ABCCT asks about the situations where unmet expectation result in negative affect. The adult version of this scale also includes a reciprocity item, as negative affect may result if one feels one is not meeting the partner's expectations.

Cost 3: Threat to Privacy. This scale includes three major aspects: concerns over your communication partner being able to see or hear something you may want to keep secret, concerns over others in the environment overhearing the communication, and concerns that the communication may be invading the partner's privacy.

Table 2. ABCCT scales and questions. <X> is the relationship and <M> is the medium being investigated. Participant should respond on a 5 point Likert-type frequency scale (never, rarely, sometimes, usually, or always)

Scale	ABCCT-Adult Questions	ABCCT-Child Questions
Emotion. Express.	Communicating with <X> using <M> helps me tell how <X> is feeling that day.	I could tell over <M> how <X> was feeling that day.
	Communicating with <X> using <M> helps me let <X> know how I am feeling.	I could let <X> know how I was feeling over <M>.
	Communicating with <X> using <M> helps me see how much <X> cares about me.	I could tell over <M> how much <X> cares about me.
Engagement & Play	-	It was boring to use <M> with <X>. [reverse score this item]
	I feel that contact with me using <M> is engaging for <X>.	-
	I am excited about using <M> with <X>.	I was excited about using <M> with <X>.
	I have fun with <X> while using <M>.	I had fun with <X> while using <M>.
Presence in-Absence	Communicating with <X> using <M> helps me feel closer to <X>.	I felt closer to <X> after using <M> together.
	After we are done communicating, I still keep thinking back to something <X> shared using <M>.	After we were done, I still kept thinking back to something <X> shared using <M>.
	Communicating with <X> using <M> helps me feel more connected to <X>.	I felt more connected to <X> after our contact using <M>.
Opportunity for Social Support	Communicating with <X> using <M> helps me provide <X> with social support.	-
	<X> makes me feel special in our contact using <M>.	<X> made me feel special in our contact using <M>.
	Communicating with me using <M> helps <X> be there for me when I need them.	<X> was there for me over <M> when I needed them.
	Communicating with <X> using <M> when I am having a bad day helps me feel better.	Using <M> with <X> when I was having a bad day helped me feel better.
	Communicating with <X> using <M> helps me feel less worried about something.	I felt less worried about something after using <M> with <X>.
Unwanted Obligations	I worry that <X> feels obligated to contact me using <M>.	-
	I have to talk to <X> using <M> even if I don't want to.	I had to talk to <X> using <M> even if I didn't want to.
	I feel guilty if I don't answer a contact <X> makes using <M>.	I felt guilty if I didn't answer a contact <X> made using <M>.
	I have to answer when <X> tries to contact me using <M> even if I don't want to.	I had to answer when <X> tried to contact me using <M> even if I didn't want to.
Unmet Expectations	I feel sad when <X> isn't around when I try to contact <X> using <M>	I felt sad when <X> wasn't around when I tried to contact <X> over <M>.
	I feel sad when <X> takes too long to respond when I try to contact <X> using <M>.	I felt sad because <X> took too long to respond when I tried to contact <X> using <M>.
	I worry that I am not meeting <X>'s expectations for our contact using <M>.	-
	I feel sad when <X> doesn't pay enough attention to me when we use <M>.	I felt sad because <X> didn't pay enough attention to me when we used <M>.
Threat to Privacy	I worry that <X> might learn something using <M> that I want to keep secret.	I worried that <X> might learn something over <M> that I want to keep secret.
	I worry about my privacy while <X> and I are using <M> together.	I worried about my privacy while <X> and I were using <M> together.
	I worry that others may overhear or see something that <X> and I share using <M>.	I worried that others would overhear or see something that <X> and I shared using <M>
	I worry that I am violating <X>'s privacy during our contact using <M>.	-

The full text of the scales and questions are presented in Table 2.

6. METHODS

We conducted three studies to gather reliability and validity measures of the ABCCT. In this section, we describe the procedures of all three studies: (1) we deployed the child version survey and discussed it in focus groups with 45 children; (2) we deployed the adult version of the survey online with 110 adult participants; (3) we conducted interviews and deployed the survey with 14 adults and 14 children in families where a parent travels frequently for work. The findings of all three studies and what these findings say

about the reliability, validity, and sensitivity of the ABCCT are related in the results section.

6.1. Deployment and Focus Groups with Children

We tested the ABCCT with a larger group of children to gather reliability statistics and continue refining the question set. The study was conducted at a school in Preston, England with 45 eight- to ten-year-old children participating in testing the survey. The survey was read aloud by the researcher to groups of 6 children as they filled it out. They were asked to complete the survey for a relationship and medium of their choosing. The majority of the children discussed using the telephone with a family member (grandparent and cousin were the most common). The chil-

Table 3. Sample interview prompts and sample coded phrases for each interview scale.

Scale	Sample Interview Prompt	Sample Coded Phrase
Emotional Expressiveness	How can you tell how your son is feeling when you talk to him on the phone?	"I have to see them to really know. But sometimes it depends on what they are saying. If they get in trouble, yeah, usually."
Engagement & Playfulness	Tell me about a time that you and your mom had fun while using telephone together.	"Usually I act pretty goofy. Like we'll say something like an inside joke, and she'll laugh. And my brother, he'll say something weird. We don't what he be saying, but we just laugh."
Presence-in-Absence	How does exchanging MMS messages affect your sense of connection to your daughter?	"I have their pictures that I look at ... if I'm in the middle of wherever, I can just pop open my cell phone and look at the pictures."
Social Support	Tell me about a time that talking to your mom over video-chat made you feel better when you were having a bad day.	"We might joke around. Like I'm feeling blue and just kind of make a face and show her ... and you know, a lot of the negative stuff, kind of like dissipates."
Feeling Obligated	Give me an example of a situation where you didn't feel like talking to your mom on IM. What did you do?	"One time, my mom responded, and I didn't really want to, so I just ignored it."
Unmet Expectations	Give me an example of an expectation that you have of your daughter in terms of responding to you on the phone.	"Yes, I expect her to return my calls, and I expect her to definitely respond to my texts or phone calls, because otherwise I get very concerned."
Threat to Privacy	Give me an example of a situation where you worried that somebody else might hear your phone conversation.	"Yes, that's why I like going home. Like, some certain things, I don't mind saying that in front of my friends. But some stuff I like to keep in the house."

dren were observed as they completed the survey and confusing questions were noted. After completing the survey, each group of six was led by the researcher to discuss which questions seemed to be most problematic. From the focus group responses, the authors identified questions that needed to be reworded or removed from the survey. The survey responses were analyzed quantitatively to gather metrics of inter-item reliability.

6.2. Deployment with Adults

To pilot the ABCCT with adults, we created an online version of the questionnaire. A link to the questionnaire was circulated using social networking media and posted in the craigslist.com "Volunteer" section. Each participant was asked to fill out the ABCCT for a relationship and communication technology of their choosing. A total of 113 participants completed the questionnaire. Two entries were removed from the analysis as the participants rated multiple communication technologies in a single survey instead of selecting one specific technology. One entry was removed because the participant gave the same answer to all questions and took less than 1 minute to complete the survey, suggesting he or she did not read the questions. The remaining 110 questionnaires were analyzed quantitatively to gather metrics of inter-item reliability and to gauge the questionnaire's sensitivity to different media.

6.3. Interviews and Surveys with Parent-Child Pairs

To continue rigorously measuring the reliability and validity of the questionnaire, we conducted a follow-up study with a smaller number of participants that allowed us to see how the results of the ABCCT corresponded to other measures of affects and benefits of communication technologies and see whether the results on the ABCCT remained stable over time for each participant.

6.3.1. Procedure

We recruited families who frequently used communication technologies and who were willing to participate in in-depth interviews. We interviewed parent-child pairs from 14 different families, where the parents and children (ages 7-13)

frequently used communication technologies to stay in touch. Recruiting was done through word-of-mouth, ads on craigslist.org, and local military support organizations. We conducted hour-long semi-structured interviews with each participant. The questions focused on the participants' general experiences with the separation, how they manage contact during separation, and how they use technology to stay in touch. We also asked about their experiences with the two most common technologies they use, specifically what they see as the costs and benefits of each technology. In addition to the interview, each parent and child responded to ABCCT questionnaires for the two technologies they used most frequently to stay in touch. We also asked each adult to fill out a Social Connectedness Questionnaire [3] specific to his or her relationship with the child interviewed (unfortunately, no child version of the Social Connectedness Questionnaire is available). Finally, all study participants were contacted again two weeks after the interview and asked to again fill out the ABCCT questionnaire for the same technologies they had previously discussed.

6.3.2. Analysis

All interviews were audio-recorded and transcribed by the lead author. The interview transcripts were analyzed by creating thematic connections using a data-driven approach [40]. With each pass through the interview data, themes were refined until a set of distinct themes emerged. In addition to this data-driven approach, for each participant, we examined all interview statements corresponding to any of the ABCCT scales. Based on these interview statements, each participant was assigned a score of either "low," "medium," "high," or "insufficient information" for each ABCCT scale. For three randomly selected participants, the lead author and an independent coder scored interview statements for each scale. The agreement (Cohen's Kappa) between the two coders was for 0.80 ("excellent") over the 42 codes. Any disagreements on codes were discussed until consensus was reached. In order to be able to quantitatively compare the results of the interview coding and the ABCCT

questionnaire responses, we also rated each participant's responses on the ABCCT according to the same scheme. If a participant averaged between 1.00 and 2.33 on a particular scale, he or she was classified as “low” on that scale; between 2.34 and 3.66 was classified as “medium” on that scale; and between 3.67 and 5.00 was classified as “high” on that scale. Thus, for each of the 14 adult and 14 child participants, for each scale of the ABCCT, we recorded two values: one as indicated by the participant through his or her ABCCT responses and one as coded by the researcher based on the participant's interview responses. We treated these two values as two sources of ratings and conducted inter-rater reliability analysis (Cohen's K) to understand the level of agreement between the two measures. Participant scales coded as “insufficient information” in the interviews were dropped from this analysis.

Lastly, we conducted a quantitative analysis of the surveys administered to these participants to get measures of convergent validity and test-retest reliability. To get a measure of convergent validity, we compared the average benefits of communication technologies as reported on the ABCCT with the Specific Social Connectedness metric of the SCQ [3]. We hypothesized that a participant that reports high level of benefits in using the group of communication technologies with a specific family member would also report having high levels of social connectedness with this family member. To get a measure of test-retest reliability, we compared two ABCCT measures taken 2 weeks apart.

6.4. Limitations

This evaluation of the ABCCT is meant to provide the first steps towards a final instrument. Though we provide several components of typical questionnaire validation, the number of participants in each part of the analysis is significantly lower than in a typical psychometrics study. For example, the PANAS affect schedule [46] was tested with an initial study of 407 participants and with a retest study of 163 (compared to 110 test and 27 retest for this ABCCT investigation). While the ABCCT approach exceeds the typical validation provided for other instruments in the HCI domain (e.g., SCQ tested with 215 participants, and provided no test-retest [3]; AttrakDif with 33 participants without test-retest [16]), it may not provide as thorough a validation as a true psychometric approach. However as a tradeoff, our work provides a metric of convergent validity with qualitative data that is rarely presented in psychometrics work but valued in HCI investigations. As HCI investigations utilize this instrument, the community may come to better understand the specific limitation it presents and which cases may require modifications of the instrument.

7. RESULTS

We present the results of the three studies described in the previous section, as related to the inter-item and test-retest reliability and convergent validity metrics of the ABCCT. We also show that the ABCCT is sensitive enough to detect differences between communication media.

Table 4. Inter-item reliability metrics for ABCCT scales.

Scale	Cronbach's α (Children)	Cronbach's α (Adult)
Benefits	0.88	0.88
Emotional Expressiveness	0.65	0.74
Engagement & Play	0.74	0.75
Presence-in-Absence	0.71	0.75
Social Support	0.86	0.75
Costs	0.80	0.85
Feeling Obligated	0.64	0.61
Unmet Expectations	0.78	0.70
Threat to Privacy	0.72	0.80

7.1. Inter-Item Reliability

We gathered inter-item reliability metrics of the ABCCT questionnaire by piloting with 45 children and 110 adults (see Table 4). Overall, the Benefits and the Costs dimensions both achieve internal consistency classified as “good.” The scales for Engagement & Playfulness, Presence-in-Absence, Opportunity for Social Support, Unmet Expectations, and Threat to Privacy all achieve Cronbach's Alphas in the “acceptable” to “good” range. However, the Emotional Expressiveness scale on the child version and the Feeling Obligated scale on both versions have a Cronbach's alpha in the “questionable” range. Since lower values are acceptable for scales with a smaller number of items [29] and one of our goals was achieving a shorter questionnaire, these scales are within reasonable parameters to warrant their use. This reflects the design decision to focus on creating a shorter questionnaire rather than focusing on the internal consistency of each individual scale. If internal consistency is of particular concern to the researcher, we recommend administering and reporting results along the entire Benefits and Costs dimensions rather than splitting the results into each specific scale.

7.2. Test-Retest Reliability

One way of demonstrating the reliability of the ABCCT is by showing that two measures of the same partner-medium interaction taken from the same participant will yield simi-

Table 5. Agreement between the interview coding and the questionnaire responses for adults and children.

Scale	Cohen's K (Adults)	Cohen's K (Children)
Overall Agreement	0.793	0.612
Emotional Expressiveness	0.85	0.63
Engagement & Play	0.78	0.29
Presence-in-Absence	0.76	0.68
Social Support	0.62	0.50
Feeling Obligated	0.71	0.60
Unmet Expectations	0.65	0.71
Threat to Privacy	0.84	0.60

lar results when separated by a fairly short period of time. In this case, it would be reasonable to expect that the measures of benefits and costs of a medium for a particular relationship would stay largely the same when taken two weeks apart. All of the adult interview participants responded to the ABCCT retest request, allowing us to compare the measures for 27 ABCCT surveys total (13 participants responded about 2 communication media each, 1 participant responded about 1 medium). For adults, the correlation (Pearson r) of the average reported benefits and costs of each communication technology on the two surveys given two weeks apart was 0.951 which is statistically significant ($p < 0.001$). This shows that for adults, the ABCCT is a stable metric over a time period of two weeks and this contributes to the argument that the ABCCT is a reliable instrument.

Gathering an ABCCT retest measure from the children required more effort on the adult participants' parts (since the questionnaire needed to be administered verbally), so not all parents responded to the request to collect these measures from their children. We collected a total of 10 ABCCT retest surveys (5 participants responded about 2 communication technologies each). However, even with the reduced sample size, the correlation (Pearson r) of the average reported benefits and costs of each communication technology on the two surveys given two weeks apart was 0.686 which is statistically significant ($p = 0.001$). This shows that for children, the ABCCT is also a stable metric over a time period of two weeks and contributes to the argument that the ABCCT is a reliable instrument.

7.3. Convergent Validity with SCQ

We asked each of the 14 adult participants to fill out a Social Connectedness Questionnaire [3] to describe their sense of connection to their child. The Social Connectedness questionnaire is neither suitable nor validated for use with children so we could only collect this metric for adults.

It would stand to reason that participants who reported receiving higher benefits from each of the communication technologies they used to maintain contact with the child remotely on the ABCCT would also report a higher level of social connectedness with the same child on the Social Connectedness Questionnaire. Indeed, the correlation

(Pearson r) between the average benefits reported on all technologies used by each adult participant and the SC score was 0.652 for the 14 adult participants, which is statistically significant ($p = 0.012$). This correlation confirms the expected relationship between the benefits dimension of the ABCCT and the construct of Social Connectedness.

7.4. Convergent Validity with Qualitative Data

Interviewing a person about the costs and benefits of a communication technology should produce roughly the same responses as asking that person to take the ABCCT. We conducted hour-long semi-structured in-depth interviews with 14 parents and 14 children. Several of the questions in the interview were explicitly designed to draw out responses related to certain ABCCT scales in an open-ended manner. Table 3 shows some of the questions that we asked in the interview to get insight into specific scales, as well as phrases from the interview that were coded as being relevant to that scale. We also coded phrases from the other parts of the interview that provided insight into specific costs and benefits of communication technologies. Since the interviewer did not pose the same questions as the survey and asked the participant to relate specific stories and experiences rather than data on frequency of particular events, not all participants addressed each scale in the interviews. If there was insufficient evidence in the interview to code a participant as “low,” “medium,” or “high” on a specific scale, that scale was dropped from the analysis (7.4% of scale-medium pairings for adults, 16.6% of scale-medium pairing for children).

On the child version, the overall agreement Cohen's Kappa was 0.612 (over 146 codes), which is statistically significant ($p < 0.001$). Table 5 shows the Cohen's Kappa for each scale on the survey. The “Engagement & Playfulness” scale stands out as having the lowest level of agreement. The interview coding almost always predicted a lower response than the one that appeared on the survey. The interview asked children to recall a time that they had fun with their parent while using a certain communication technology together. Many of the children could not recall a single instance of “having fun.” Yet, on the survey the same children reported that talking to their parent was “never” boring and “sometimes” or “frequently” fun. There could be two possible reasons for this observation. First, it's possible that

Table 6. Comparing different communication technologies using the ABCCT, in the adult pilot deployment. All comparisons were done using 2-tailed Wilcoxon rank sum tests. Values in square brackets are not statistically significant, but show a trend.

Scale	Phone (n=43) vs. Video (n=20)			IM (n=19) vs. Text (n=16)			Video (n=20) vs. Text (n=16)		
	Phone	Video	M.-Wh.	IM	Text	M.-Wh.	Video	Text	M.-Wh.
Emotional Expressiveness	3.86	4.02	not sig.	3.67	3.15	$p = 0.023$	4.02	3.15	$p < 0.001$
Engagement & Play	3.42	3.84	$p = 0.044$	3.54	3.14	$p = 0.047$	3.84	3.14	$p = 0.004$
Presence-in-Absence	3.84	4.30	$p = 0.027$	3.51	3.17	not sig.	4.30	3.17	$p < 0.001$
Social Support	3.57	3.68	not sig.	3.54	3.50	not sig.	3.68	3.50	not sig.
Feeling Obligated	2.22	1.79	$p = 0.049$	2.32	2.03	$[p = 0.160]$	1.79	2.03	$[p = 0.170]$
Unmet Expectations	2.36	2.14	not sig.	2.51	2.22	not sig.	2.14	2.22	not sig.
Threat to Privacy	1.81	2.04	not sig.	1.89	1.39	$p = 0.013$	2.04	1.39	$p = 0.038$

the children were simply not very capable of recalling specific past examples, but may in fact have had a better gestalt impression of their engagement with their parents. Alternatively, it is possible that children experienced greater demand characteristics with the survey—a child may not want to say that talking with their parent was boring for fear of hurting their parents' feelings. In this situation, it is difficult to know whether the interview or the survey response is closer to the ground truth. With children, it seems that one may have to triangulate evidence from multiple sources in order to get an accurate understanding of their engagement. However, it seems that on the other scales, the interview and the ABCCT responses corresponded significantly for each participant—evidence that the ABCCT-Child measures what it intends to measure.

On the adult version, the overall agreement Cohen's Kappa was 0.793 (over 175 codes), which is statistically significant ($p < 0.001$). Table 5 shows the Cohen's Kappa for each scale on the survey. The adult version of the ABCCT showed significant agreement on all scales between the experiences related in the interview and the responses given on the written ABCCT. These results provide strong evidence that the ABCCT-Adult measures the same construct as gleaned from independently coded interviews.

7.5. Sensitivity to Differences between Technologies

One of our goals in this undertaking was creating a questionnaire that is sensitive enough to detect differences in emotional benefits and costs between different communication media. We used the data from the ABCCT administered online to 110 adults to address this concern. Unfortunately, we could not conduct a similar analysis for ABCCT-Child because there were fewer than 5 responses addressing any technology other than the telephone in the Preston pilot, so we could not make a meaningful quantitative comparison. Since the adult pilot gathered a number of responses for a fairly diverse set of communication technologies, we could conduct three comparisons. First, we compared two synchronous and two asynchronous communication technologies most frequently described in the online surveys.

We used a two-tailed Wilcoxon Rank Sum Test to compare telephone vs. videoconferencing, IM vs. texting, and texting vs. videoconferencing on each of the 7 ABCCT scales. Table 6 shows the results of these comparisons. As could be expected, videoconferencing scored significantly higher on the Engagement & Playfulness and Presence-in-Absence scales. This is consistent with previous qualitative findings in the field (e.g., [1,2]). Participants also felt more unwanted obligations with the telephone. Though there was no statistically significant difference between telephone and videoconferencing on Emotional Expressiveness and Threat to Privacy scales, both were trending in expected directions. Comparing IM and Texting showed that IM scored higher on measures of Emotional Expressiveness and Engagement & Playfulness. Interview data suggests that this might be due to IM's support for images, emoticons, and ability to

link to other media (e.g., YouTube clip) [52]. There is a trend (through not statistically-significant) for IM introducing more unwanted obligations, which is consistent with Grinter et al.'s qualitative work examining IM and SMS [14]. Also as expected, IM introduced significantly more threats to privacy than texting, due to the fact that IM conversations can easily be saved and passed along to others. Lastly, we compared videoconferencing and texting, because these two media have been previously compared on social presence measures in the lab [15]. These two media are quite different and we expected to see differences on a number of scales. Indeed, we found that videoconferencing was significantly better at supporting emotional expression, engagement & playfulness, and presence-in-absence. Also, as can be expected, videoconferencing introduces significantly greater concerns over privacy. Texting showed a non-statistically significant trend for more obligations to communicate than videoconferencing, which is consistent with interview findings that users are often expected to answer their texts in a timely manner (e.g., [52]). It is likely that there was significant noise in this data, since the survey was administered online with very little guidance for participants. For example, participants may have used the same medium label (e.g., "IM") for two different experiences (e.g., Facebook messenger vs. Pidgin). It is encouraging that the ABCCT was able to highlight some statistically significant differences between communication technologies even with this significant noise, a relatively low number of responses, and data collected between (rather than within) participants.

8. EARLY QUESTIONNAIRE UPTAKE

The success of the ABCCT depends on community uptake. Though this paper presents the first archival publication of the ABCCT validation, we made early versions of the questionnaire available to other researchers. There are three independent examples of early uptake of this questionnaire prior to this paper. One project used the ABCCT to understand the ecosystem of communication tools available to users for sharing information and support in dealing with chronic health conditions [31]. The ABCCT helped these researchers articulate a problem space for novel technology and the different roles that synchronous and asynchronous communication media play in health-related communication. Another group of researchers, incorporated the ABCCT into their methodology to understand the role that telepresence robot can play in connecting separated family members [48]. In the third project, the ABCCT was used to evaluate a novel communication technology for connecting separated parents and children. The ABCCT was able to demonstrate the specific scales where the novel technology performed better than the previous communication media used by the families in an A-B field deployment [56]. These initial forays may be representative of the community's interest in incorporating validated questionnaires as one of the sources of evidence in understanding the emotional costs and benefits of communication technologies.

9. DISCUSSION

We provide guidelines for scoring and use of the ABCCT, discuss its limitations, and offer directions for future work.

9.1. ABCCT Questionnaire Scoring and Analysis

The ABCCT uses Likert-type frequency scales with the assumption of equal interval between items. The investigator should assign numbers from 1 to 5 for each item (i.e., “never” = 1, “always” = 5). One question on the child version (marked in Table 2) is to be scored negatively (i.e., assign score of 6 - response). Once scores for each question are obtained, the score for each scale can be computed by averaging those responses. The score for each dimension can be calculated by averaging the scores for each composing scale. With descriptive investigations, scores on each scale and/or dimension 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 paired (within subjects design) and non-paired (between subject design) data. 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.

9.2. Guidelines for ABCCT Questionnaire Use

The ABCCT can be administered to adults on paper or online. The pilot deployment with 110 adults was completed online with minimal instruction for participants. In administering the child version of the questionnaire, we recommend following Borgers & Hox's guidelines [29]: the questionnaire should be clearly explained, the questions should refer to the opinions and behaviors of a recent time period (e.g., past 2 weeks), and all options on the survey should be labeled (e.g., “Never,” “Rarely,” etc.). When possible, we recommend that the survey be administered as a structured interview to children (i.e., read out loud to the child), as this helps ameliorate individual differences in reading level and comprehension. Additionally, this allows the researcher to verbally make the substitutions for the “person” and “medium” being investigated.

The main value of the ABCCT is in allowing the investigator to make comparisons between different technologies. Here are several potential scenarios for appropriate use of the ABCCT:

- **Within-Subject Comparison of Two Tech Versions:** Two versions of a communication technology (with and without an experimental feature) are deployed one-after-another in a series of homes. Each participant fills out the ABCCT for each technology after the deployment. A pairwise comparison is used to investigate how the feature affected the benefits and costs of the technology.
- **Between-Subject Comparison of Two Tech Versions:** One version of a communication technology (either with or without an experimental feature) is deployed in the home. After the deployment, each participant fills out the

ABCCT for the technology version they used. Scores on each scale or dimension are compared between the control and experimental groups to investigate how the feature affected the benefits and costs of the technology.

- **Within-Subject Comparison of Prototype Tech with Existing Tech:** A technology is deployed to replace a particular current technology (e.g., better IM client). Before the study, the participant is asked to fill out the ABCCT for the current technology. After the deployments, the participant is asked to fill out the ABCCT for the novel system. A pairwise comparison is used to point out how the novel system differs from the status quo system (e.g., [56]).
- **Within-Subject Pre-Post Deployment Tech Ecosystem Comparison:** A technology that is meant to supplement the currently used communication systems is deployed in the field. Before and after the deployment, the participants fill out ABCCT questionnaires for each of the communication media they use. An analysis can show whether the new system filled a need different from the other systems, whether previous systems were now used differently because of the introduction of the novel technology, and how the novel system compares to the previously used systems on each of the ABCCT dimensions.
- **Meta-Analysis Regarding a Particular Feature:** An investigator wants to conduct a meta-analysis showing that adding a particular feature (e.g., asynchronous video messaging) to a communication system has a predicted effect on a specific ABCCT scale. He or she gathers all papers that investigated relevant systems with this feature and all papers that investigated relevant systems without this feature. Assuming these papers presented ABCCT scores (or made these metrics available by request), the investigator can use standard meta-analysis techniques to draw comparisons and conclusions.
- **Meta-Analysis Regarding a Particular Population:** An investigator wants to conduct a meta-analysis of ABCCT scores for systems supporting a particular population (e.g., remote couples). He or she gathers all papers that included this population. Assuming these papers presented ABCCT scores (or made these metrics available by request), 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.

While this is not an exhaustive list of acceptable study designs, it provides some examples of how the ABCCT may be used. We caution investigators against making between-study comparisons when only a small number of other studies are available on a particular issue of interest. Particularly, **drawing favorable or unfavorable comparisons against a technology in one previous investigation is more likely to be misleading than informative.** Brown et al. point out the incredible difficulties in attempting to generalize from single-case field deployments, as even small variations in participants and methods can lead to drastical-

ly different interpretations of the same technology [5]. However, as similar technologies are deployed in a variety of contexts, there may be a large enough sample to compensate for this variability and allow for between-study comparisons.

9.3. Limitations

Researchers should understand the inherent limitations of the ABCCT questionnaire before using it. Most importantly, **this questionnaire should not be used as the only evaluation metric for a novel communication technology**. The emotional aspects of communication are inherently nuanced. While the ABCCT provides a good way to compare systems and synthesize findings across investigations, qualitative methods are still the best way of understanding the contextual factors that play into the adoption and use of a novel system for communication. However, triangulating data from rich qualitative metrics and the structured data from the ABCCT would lead to more valuable results than using either method alone.

Second, **the adoption and use of communication technologies is influenced not just by emotional factors but also by pragmatic time/effort-related factors and even basic usability issues**. For example, videochat is frequently not used by families because of the complex work required to set up a session [1] and because of the undivided attention demanded by this medium [52]. On the other hand, communication media that have low time and effort costs to compose and send a message may be compelling even if they do not provide a rich channel for communication (e.g., [10]). Measuring the costs of initiating contact, the costs of composing or maintaining the contact, and the opportunity provided for demonstrating personal effort are all meaningful metrics for communication technologies but outside of the scope of the ABCCT. To gain a more complete understanding of the technology being investigated, the **ABCCT may be combined with other validated metrics** (e.g., usability metrics).

Lastly, the ABCCT questionnaire has the same limitations as other questionnaire methods appropriate for field evaluations. The participants fill out the ABCCT some time after events of interests occur, so **(1) participants may forget important issues or have trouble estimating the frequency of particular events. (2) Participants may be biased to present themselves in a positive light to the investigator** and therefore choose responses that may not accurately reflect their experiences with the technology. Additionally, **(3) participants may respond to the demand characteristics of the investigator**, especially when the expected results easy to deduce (e.g., within-subjects design where the experimental system is obvious). This may be especially true of children, who often worry about providing the adult with “the right answer.” Triangulating the results from multiple methods and multiple sources of data can mitigate these shortcomings. Additionally, taking care to equalize power between the child and the researcher

using classic participatory design methods can address some inequalities that exacerbate demand characteristics with children (e.g., [25]).

9.4. Future Directions

There are a number of future directions that emerge from this work: (1) designing a methodology for measuring the “pragmatic” costs and benefits of a communication technology, (2) using the ABCCT in the field and iterating on the design when necessary, (3) creating alternative versions of the ABCCT that may be useful for experience sampling methodologies.

The ABCCT does not provide a complete picture of a participant's experience with a communication technology. Understanding the pragmatic costs and benefits of communication technologies, such as the time and effort expended during use, may help complete this picture. It may be possible to gather some of this data using survey methodologies, but it may be impractical due to the difficulty that users have recalling the exact amount of time spent on a particular interaction. Logging methodologies may be most valuable in this domain, perhaps in combination with some appropriate self-response data (for example, getting a participant to reflect on specific log data in an interview).

One of the potential benefits of the ABCCT is in supporting synthesis and comparison across technologies in multiple studies. However, in order for this to be possible, investigators must include the ABCCT in their methods and report (or make available upon request) the results of these deployments. As the ABCCT is utilized in diverse contexts, it is quite likely that shortcomings will emerge that will need to be addressed in future iterations. As such, we greatly encourage any investigators who chose to use the ABCCT to share their experiences with the authors of this instrument. We also encourage investigators to modify the ABCCT to suit the needs of a particular research project, though with the caution that some modification may need to be re-validated before use.

One of the limitations of the ABCCT is that it asks participants to reflect upon past experiences, potentially leading to recall bias. Ideally, the survey could be adopted to pose appropriate questions to the participants immediately after an event of interest has occurred. For example, immediately after receiving a message from a loved one or terminating a synchronous communication session, the participant can be asked to answer 2–3 questions about their experience. While it would be impractical to administer the entire ABCCT after each session with the technology, a system may be designed to ask a random subset of relevant questions. Collecting such responses over the course of a long deployment may contribute to an understanding of the system with less recall bias. Such a method would not only involve rewording the questions to be relevant in the moment, but also designing the system that would provide the participant with the questions at the appropriate time.

10. CONCLUSION

This work presents the design and initial validation of a questionnaire instrument for measuring the Affective Benefits and Costs of Communication Technologies (ABCCT). We discuss the pilots of this survey with 45 children and 110 adults, leading to the preliminary conclusion that the ABCCT has acceptable inter-item reliability. We discuss the results of interviews with 14 children and 14 adults, which help confirm that the ABCCT measures the same constructs that may emerge through an interview investigation. We show that the ABCCT has high test-retest reliability and convergent validity with the Social Connectedness metric. Finally, we demonstrate that the ABCCT is sensitive enough to discriminate between different communication technologies and has shown promise in some of its early uptake. Though the ABCCT is not without limitations, it provides a step towards a reliable and valid way to compare communication technologies in field deployments.

ACKNOWLEDGMENTS

We would like to acknowledge Janet Read for helping us set up the Preston trial. We would also like to thank the many participants who shared their time and honesty with us during this project.

REFERENCES

1. Ames, M.G., Go, J., Kaye, J., and Spasojevic, M. Making Love in the Network Closet: The Benefits and Work of Family Videochat. *Proc. of CSCW*, (2010), 145–154.
2. Ballagas, R., Kaye, J. “Jofish,” Ames, M., Go, J., and Raffle, H. Family communication: phone conversations with children. *Proc. of IDC*, ACM (2009), 321–324.
3. Bel, D.T. van, Smolders, K.C.H.J., IJsselsteijn, W.A., and Kort, Y.A.W. de. Social connectedness: concept and measurement. *International Conference on Intelligent Environments*, IOS Press (2011), 67–74.
4. Bell, A. Designing and testing questionnaires for children. *Journal of Research in Nursing* 12, 5 (2007), 461–469.
5. Bonner, J.V.H. Adding critical sensibilities to domestic communication technologies. *Intl J of Human Computer Studies* 67, 2 (2009), 215–221.
6. Borgers, N. and Hox, J.J. Reliability of Responses in Questionnaire Research with Children. *International Conference on Logic and Methodology*, (2000).
7. Brown, B., Reeves, S., and Sherwood, S. Into the Wild: Challenges and Opportunities for Field Trial Methods. *Proc. of CHI*, (2011), 1657–1666.
8. Caughlin, J.P. and Petronio, S. Privacy in Families. In *Handbook of Family Communication*. Lawrence Erlbaum Associates, Mahwah, NJ, 2004, 379–412.
9. Dalsgaard, T., Skov, M.B., Stougaard, M., and Thomassen, B. Mediated intimacy in families: understanding the relation between children and parents. *Proc. of IDC*, ACM (2006), 145–152.
10. Dalsgaard, T., Skov, M.B., and Thomassen, B.R. eKiss: Sharing Experiences in Families Through a Picture Blog. *Proc. of British HCI*, (2007), 67–75.
11. Feltham, F., Vetere, F., and Wensveen, S. Designing tangible artefacts for playful interactions and dialogues. *Designing Pleasurable Products and Interfaces*, ACM (2007), 61–75.
12. Fitness, J. and Duffield, J. Emotion and Communication in Families. In *Handbook of Family Communication*. Lawrence Erlbaum Associates, Mahwah, NJ, 2004, 473–494.
13. Gardner, K.A. and Cutrona, C.E. Social Support Communication in Families. In *Handbook of Family Communication*. Lawrence Erlbaum Associates, Mahwah, NJ, 2004, 495–512.
14. Grinter, R.E., Palen, L., and Eldridge, M. Chatting with teenagers: Considering the place of chat technologies in teen life. *ACM ToCHI* 13, 4 (2006), 423–447.
15. Harms, C. and Biocca, F. Internal consistency and reliability of the networked minds social presence measure. (2004), 246–246.
16. Hassenzähl, M., Burmester, M., and Koller, F. At-trakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. *Mensch & Computer: Interaktion in Bewegung*, (2003), 187–196.
17. Hindus, D., Mainwaring, S.D., Leduc, N., Hagström, A.E., and Bayley, O. Casablanca: designing social communication devices for the home. *Proc. of CHI*, (2001), 325–332.
18. Howard, S., Kjeldskov, J., Skov, M., Garnoes, K., and Grünberger, O. Negotiating presence-in-absence: contact, content and context. *Proc. of CHI*, (2006), 909–912.
19. IJsselsteijn, W., Van Baren, J., Markopoulos, P., Romero, N., and de Ruyter, B. Measuring Affective Benefits and Costs of Mediated Awareness: Development and Validation of the ABC-Questionnaire. In *Awareness Systems*. 2009, 473–488.
20. Kaye, J. “Jofish,” Levitt, M.K., Nevins, J., Golden, J., and Schmidt, V. Communicating intimacy one bit at a time. *Proc. of CHI*, ACM Press (2005), 1529–1529.
21. Kennedy, T.L.M., Smith, A., Wells, A.T., and Wellman, B. “Networked Families.” *Pew Internet & American Life Project Report*. Washington, D.C., 2008.
22. Khan, J.V., Markopoulos, P., Eggen, B., and Metaxas, G. Evaluation of a pervasive awareness system designed for busy parents. *Pervasive and Mobile Computing* 6, 5 (2010), 537–558.
23. Khan, V.-J. and Markopoulos, P. Busy families’ awareness needs. *Intl J of Human Computer Studies* 67, 2 (2009), 139–153.
24. Lindley, S.E., Harper, R., and Sellen, A. Designing a technological playground: a field study of the emergence of play in household messaging. *Proc. of CHI*, ACM (2010), 2351–2360.
25. Markopoulos, P., Read, J.C., MacFarlane, S., and Hoysniemi, J. *Evaluating Children’s Interactive Products: Principles and Practices for Interaction Designers*. Morgan Kaufmann, Amsterdam, 2008.

26. Mynatt, E., Rowan, J., Jacobs, A., and Craighill, S. Digital family portraits: supporting peace of mind for extended family members. *Proc. of CHI*, ACM Press (2001), 333–340.
27. Neustaedter, C., Brush, A., and McDonald, D.W. Designing for Families. *Ext. Abst. of CSCW*, (2008).
28. Noller, P. and Feeney, J.A. Studying Family Communication: multiple methods and multiple sources. In *Handbook of Family Communication*. Lawrence Erlbaum Associates, Mahwah, NJ, 2004, 31–50.
29. Nunnally, J. and Bernstein, I. *Psychometric Theory*. McGraw-Hill, 1994.
30. Olivier, P. and Wallace, J. Digital technologies and the emotional family. *Intl J of Human Computer Studies* 67, 2 (2009), 204–214.
31. Pang, C.E., Neustaedter, C., Riecke, B.E., Oduor, E., and Hillman, S. Technology preferences and routines for sharing health information during the treatment of a chronic illness. *Proc. of CHI*, (2013), 1759–1768.
32. Petersen, M.G. Squeeze: designing for playful experiences among co-located people in homes. *Ext. Abst. of CHI*, ACM (2007), 2609–2614.
33. Pierce, G.R. Quality of Relationships Inventory: Assessing the Interpersonal Context of Social Support. *Commun. of Social Support*, (1991), 246–266.
34. Raffle, H., Ballagas, R., Revelle, G., et al. Family Story Play: Reading with Young Children (and Elmo) Over a Distance. *Proc. of CHI*, ACM (2010).
35. Raffle, H., Revelle, G., Mori, K., et al. Hello, is Grandma there? Let's Read! StoryVisit: Family Video Chat and Connected E-Books. *Proc. of CHI*, (2011), 1195–1204.
36. Romero, N., van Baren, J., Markopoulos, P., de Ruyter, B., and IJsselsteijn, W. Addressing Interpersonal Communication Needs through Ubiquitous Connectivity: Home and Away. *Ambient Intelligence*, (2003), 419–429.
37. Romero, N., Markopoulos, P., van Baren, J., de Ruyter, B., IJsselsteijn, W., and Farshchian, B. Connecting the family with awareness systems. *Personal and Ubiquitous Computing* 11, 4 (2007), 299–312.
38. Rowan, J. and Mynatt, E.D. Digital Family Portrait Field Trial: Support for Aging in Place. *Proc. of CHI*, ACM (2005), 521–530.
39. Saslis-Lagoudakis, G., Cheverst, K., Dix, A., Fitton, D., and Rouncefield, M. Hermes@Home: supporting awareness and intimacy between distant family members. *Proc. of OZCHI*, ACM (2006), 23–30.
40. Seidman, I. *Interviewing as Qualitative Research: A Guide for Researchers in Education And the Social Sciences*. Teachers College Press, New York, 1998.
41. Sellen, A., Harper, R., Eardley, R., et al. HomeNote: Supporting Situated Messaging in the Home. *Proc. of CSCW*, (2006), 383–392.
42. Short, J., Williams, E., and Christie, B. *The social psychology of telecommunications*. Wiley, London, 1976.
43. Strong, R. and Gaver, W.W. Feather, Scent, and Shaker: Supporting Simple Intimacy. *Proc. of CSCW*, (1996).
44. Tee, K., Brush, A.J., and Inkpen, K.M. Exploring communication and sharing between extended families. *Intl J of Human Computer Studies* 67, 2 (2009), 128–138.
45. Teh, J.K.S., Cheok, A.D., Peiris, R.L., Choi, Y., Thung, V., and Lai, S. Huggy Pajama: a mobile parent and child hugging communication system. *Proc. of IDC*, (2008), 250–257.
46. Thompson, E.R. Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology* 38, 2 (2007), 227–242.
47. Tollmar, K., Junestrand, S., Torgny, O., Boyarski, D., and Kellogg, W.A. Virtually living together. *Proc. of DIS*, (2000), 83–91.
48. Tsui, K., Desai, M., and Yanco, H. Towards measuring the quality of interaction: Communication through telepresence robots. *Proc. Workshop on Performance Metrics for Intelligent Systems*, ACM (2012).
49. Vetere, F., Davis, H., Gibbs, M., and Howard, S. The Magic Box and Collage: Responding to the challenge of distributed intergenerational play. *Intl J of Human Computer Studies* 67, 2 (2009), 165–178.
50. Vetere, F., Gibbs, M.R., Kjeldskov, J., et al. Mediating intimacy: designing technologies to support strong-tie relationships. *Proc. of CHI*, ACM (2005), 471–480.
51. Visser, T., Dadlani, P., van Bel, D., and Yarosh, S. Designing and Evaluating Affective Aspects of Sociable Media to Support Social Connectedness. *Ext. Abst. of CHI*, (2010), 4437–4440.
52. Yarosh, S. and Abowd, G.D. Mediated Parent-Child Contact in Work-Separated Families. *Proc. of CHI*, ACM (2011), 1185–1194.
53. Yarosh, S., Chew, Y.C., and Abowd, G.D. Supporting Parent-Child Communication in Divorced Families. *Intl J of Human Computer Studies* 67, 2 (2009), 192–203.
54. Yarosh, S., Inkpen, K.M., and Brush, A. Video playdate: toward free play across distance. *Proc. of CHI*, (2010), 1251–1260.
55. Yarosh, S. and Markopoulos, P. Design of an instrument for the evaluation of communication technologies with children. *Proc. of IDC*, ACM (2010), 266–269.
56. Yarosh, S., Tang, A., Mokashi, S., and Abowd, G. “Almost Touching:” Parent - Child Remote Communication Using the ShareTable System. (2013), 181–192.