Feasibility and acceptability of a novel tool for the study of interpersonal processes in psychotherapy

Carly Max Schwartzman
University at Albany, State University of New York, carly.schwartzman@gmail.com

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Clinical Psychology Commons

Recommended Citation

This Master's Thesis is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive. Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
FEASIBILITY AND ACCEPTABILITY OF A NOVEL TOOL FOR THE STUDY OF
INTERPERSONAL PROCESSES IN PSYCHOTHERAPY

by

Carly M. Schwartzman

A Thesis
Submitted to the University at Albany, State University of New York
in Partial Fulfillment of
the Requirements for the Degree of
Master of Science

College of Arts & Sciences
Department of Psychology
2018
Abstract

Psychotherapy process research methods often require extensive time and resources. Technology innovations have the potential to increase the efficiency of data collection and processing. A technology with potential applications for psychotherapy research is the Sociometric Badge (SB), which is a portable, palm-sized device that can simultaneously record session audio and data on social signals (e.g., speech patterns, body movement) in real-time and in varied contexts. This pilot study examined the feasibility and acceptance of these assessment devices in comparison with traditional audio recording equipment. Undergraduate students (N = 308; Mage = 19.16 years [SD = 1.4]; 50.3% female) were randomly placed into 154 dyads, in order to mimic a psychotherapy dyad. Each dyad was randomly assigned to either a SB condition (n = 75 dyads) or a standard recording device condition (n = 79 dyads), and was instructed to engage in distinct conversation tasks. At the end of the interactions, participants completed self-report items that assessed their perceived relationship quality with their partner and experience of the respective recording device. Between-condition tests were performed with HLM and MANOVA. As hypothesized, study condition was not a significant predictor of perceived relationship quality. Similarly, device satisfaction generally did not differ between conditions. However, participants in the audio recorder condition reported more awareness of the device than those in the SB condition. These findings reveal comparable acceptability and feasibility of SBs to traditional audio recorders in a simulated psychotherapeutic dialogue. Results suggest that these devices may be suitable for research in an authentic psychotherapy setting.

Keywords: Sociometric Badge; technology; psychotherapy process; research methods
Introduction

Psychotherapy process research is interested in the study of variables that span participant (patient and therapist) characteristics, technical factors (e.g., therapist providing interpretations), and relationship variables (e.g., working alliance). Depending on the context, any of the above categories of variables could be a dependent variable of interest; however, researchers are often interested in the associations between process and outcome, such as immediate or long-term changes that occur as a result of the therapy (e.g., decreases in client symptoms). The main goal of this research is to identify, describe, and predict the processes that bring about therapeutic change (Crits-Christoph, Connolly Gibbons, & Mukherjee, 2013). The general effectiveness of bona fide psychotherapies is well-established (Nathan & Gorman, 2015), yet even within the context of thoroughly researched treatments, precise mechanisms of change are still largely unknown (Crits-Christoph et al., 2013; Kazdin, 2007). Furthermore, a notable percentage of patients do not respond to a standard course of treatment (Crits-Christoph et al., 2013).

Process research is a necessary complement to randomized clinical trials (RCTs) and other forms of efficacy research, in order to gain a better understanding of how psychotherapy works, for whom, and under which circumstances (Elliott, 2010). Perhaps out of necessity, process research is broad and diverse in its focus and methodologies. The present study reports the results from a preliminary investigation of a novel assessment tool that has promise for studying psychotherapy process and outcome, and thereby, increasing the field’s knowledge about mechanisms of change. Before describing this tool and the present study in more detail, the following section briefly summarizes key trends in psychotherapy research. An understanding of these trends will highlight the potential utility of the assessment tool under investigation.
Trends in Process Research

A common focus of psychotherapy research has been the identification of participant characteristics that help explain variability in outcome. By far, most of this work has focused on demographic variables (e.g., age, gender, ethnicity), likely because these are “cheaply” obtained. Although some participant variables, such as baseline severity (Elkin et al., 1995; Thase et al., 1997), have been associated with treatment outcome, the size of demographic individual difference effects when present tends to be weak (Boswell, Constantino, & Anderson, 2016). A more recent example of this general observation can be found in the therapist effects literature. The therapist effect is used to describe the findings that some therapists are observed (statistically) to achieve consistently superior performance (e.g., outcomes, alliance, and fidelity), while others are observed to consistently achieve poorer performance (Castonguay & Hill, 2017). A logical question is whether this effect can be explained by characteristics of the therapist (e.g., more effective therapists have more experience).

Anderson and colleagues (2009) explain that generally, years of experience and demographic characteristics such as therapist age, sex, and ethnicity have failed to emerge as predictors of outcome (Beutler et al., 2004) or between-therapist differences (Constantino, Boswell, Coyne, Kraus, & Castonguay, 2017). Okiishi, Lambert, Nielsen, and Ogles (2003) examined the outcomes of 1,841 clients seen in a university counseling center and found significant variability among therapists in their clients’ outcomes. In further analyses on a larger version of that dataset, Okiishi et al. (2006) found that therapist effects were not attributable to a variety of therapist traits, such as sex, type of training, or theoretical orientation. Results from other studies have shown that significant between-therapist differences in outcome can be observed even in controlled research with highly and uniformly trained therapists strictly
adhering to treatment manuals (for a review see Baldwin & Imel, 2013). Overall, therapist effects research, both naturalistically and with clinical trials data, seems to support the hypothesis that individual therapists differentially impact client outcomes (Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009). However, very little is known about what accounts for these effects (Castonguay & Hill, 2017).

A candidate “mechanism” of the therapist effect is the ability to form a positive working alliance. Baldwin, Wampold, and Imel (2007) demonstrated that between-therapist outcome differences were significantly, yet still only partially, accounted for by between-therapist differences in working alliance scores. Therapists with better alliances on average tended to be better performing therapists on average. This might seem intuitive; however, its importance is highlighted by the lack of observed between-patient (within-therapist) associations with outcome in the same study, meaning that the effects were truly being driven by therapist differences rather than client differences within a therapist’s caseload.

More broadly, the working alliance and other relationship variables have received significant attention in the process research literature. Extensive research demonstrates the impact of the patient-therapist relationship on therapy outcomes (e.g., Goldfried, 2013; Horvath, Del Re, Flückiger, & Symonds, 2011). A positive relationship is often associated with successful outcomes, while strains in the relationship are often associated with therapeutic failure (e.g., Thomas, Hopwood, Woody, Ethier, & Sadler, 2014). Although the universality of a strong and consistent alliance-outcome association is perhaps overstated, the importance of the working alliance, in particular, is generally accepted (Castonguay, Constantino, Boswell, & Kraus, 2010). More problematically, some researchers have debated the relative importance of the working alliance and other so-called common factors versus technical factors (Castonguay, Constantino,
& Holtforth, 2006; Iwakabe, Rogan, & Stalikas, 2000). This “zero sum game” approach is at the very least an oversimplification of the treatment process. The working relationship, for example, provides the context in which interventions are implemented, and the misapplication of interventions has the potential to lead to alliance ruptures (Safran, Muran, & Eubanks-Carter, 2011; Safran, Crocker, McMain, & Murray, 1990).

Importantly, most of the research highlighted above has relied on participant self-report. Observational measures of the alliance and interpersonal process more broadly have been developed and possess adequate psychometric properties, yet they are utilized less frequently than patient self-report (Horvath et al., 2011). Conversely, the gold standard approach to fidelity (adherence and competence) monitoring is independent evaluation and ratings of recorded sessions. These trends are relevant for several related reasons. First, even when observing recorded sessions, most of these ratings (e.g., alliance quality and competence) are conducted and analyzed at the session level—each session receives an alliance score and an adherence score, so little is known about the dynamic nature of the interaction as it unfolded during the session. Second, human observational assessment is more resource intensive than self-report administration. Human coders must be trained to reliability and spend time reviewing the recordings. Consequently, the gold standard approach is almost exclusively used in funded controlled research. Third, even when there are resources to support independent coding, researchers are forced to, somewhat arbitrarily, sample specific sessions (or even random segments of sessions) and specific patients due to time and other resource limitations.

To summarize, more interpersonally oriented and dynamic process variables (e.g., alliance, competence) appear to be more consistently predictive of outcome than static variables such as participant demographic characteristics. Because of their dynamic nature, assessment of
within-session process, although not always a requirement, is the ideal approach. Studying the dynamic aspects of the treatment process—how relationship and other dyadic variables develop, vary, and change—is critically important for understanding effective therapy (Henry, Schact, & Strupp, 1986; Thomas et al., 2014). A potentially promising approach for identifying the sources of therapist effects, for example, would be to operationalize and examine therapists’ skills in facilitating therapy processes (e.g., empathy and the alliance) that are theoretically and empirically related to therapy outcome (Anderson et al., 2009). This perspective has led to the investigation of interpersonal transactions in the therapy dyad as the fundamental unit of psychotherapy process analysis. However, the variables measured and measurement tools developed for this research have significant limitations that will be further explored below (Comer & Kendall, 2013).

**Process Research Methods and Limitations**

The most frequently used measurement tools in process research are interval scales, nominal category systems, and Q sorts (Comer & Kendall, 2013). However, many process variables do not fit the necessary assumptions for interval data, in that there is not an equal difference between points on the scale. Respondents typically make judgments about therapy sessions from different stimuli, such as live sessions, session transcripts, audiotapes, videotapes, or some combination of these media. So-called micro-process measures examine moment-by-moment experiences during therapy in units such as single words, phrases, sentences, speaking turns, and thought units.

Micro-analytic sequential process research focuses on coding client and therapist responses on a relatively small number of categories or rating scales (Elliot et al., 2010). Such studies examine the direct, immediate influence of therapeutic interventions on within-session
client processes and the effect of client actions on the processing and planning activities of the therapist. As this methodology closely follows the tangible actions of client and therapist, it has potential for testing claims about causality of therapeutic influence (Elliott et al., 2010). Additionally, results of these studies may be useful in guiding clinical practice because they measure very concrete therapy processes.

However, the process research required to examine potential mechanisms of change at these levels involves very labor intensive, time consuming, and cost-prohibitive protocols, barring many researchers from conducting large-scale studies. For example, most measures used to code within session processes require human judges, which potentially introduces noise and undependability in the data (Comer & Kendall, 2013). Ultimately, judges interpret data on the basis of their own reactions, which are often accompanied by potential implicit and explicit bias. To reduce some bias, the judges must first be trained on the coding procedure. Judges should preferably be trained by the developer of the measure of choice, or by individuals with demonstrated proficiency in using the measure to ensure it is used as intended. However, this type of training is not always available, and when available may take several months to complete. In addition, after completing training, judges must work to attain adequate inter-rater reliability. Yet, many variables of clinical interest require inference, often resulting in low reliability (Comer & Kendall, 2013). Judges may need several weeks to months of practice post-training to reach acceptable reliability standards. Once judges have attained reliability, coding the study data requires several more months to years of work as well.

Best practice also often entails using multiple judges to generate credible data, which can become expensive and is sometimes impractical. Furthermore, rater drift—the phenomenon in which judgments become less reliable as familiarity increases and routine becomes established—
can cause problems in the data and must be carefully watched. Bias can also occur when judges are unable to make objective judgments, perhaps because of their current mood, attitude, or personality. Although graduate students are often considered a valuable resource as a “working force” in coding, there have been controversies about the capacity of clinically inexperienced students to evaluate complex clinical phenomena (Schanche, Nielsen, McCullough, Valen, & Mykletun, 2010).

Another problem area arises in determining the acceptable number of therapy sessions or segments within sessions to code to allow for generalization of results to the entire treatment sample. Although examining the whole course of therapy would be ideal, the time-consuming nature of coding most likely deems this financially and practically infeasible. Consequently, researchers must sample segments of the data. One specific problem in terms of sampling involves infrequently occurring events such as self-disclosure, crying, or laughter that may be missed if only certain segments or sessions are coded. Researchers may need to review a large number of sessions from a variety of therapists to find enough instances of these events, which again is often infeasible. For similar reasons, many studies using micro-analytic process research methods are contained to employing small samples, which decreases statistical power and limits the conclusions that can be drawn. Other global limitations of the literature include differing definitions of the same construct, lack of replication or inconsistency of results when such attempts are made, correlational data that limit conclusions about causality, and a focus on one type of therapy without comparable analyses for a control therapy, restricting comparisons across studies (Garfield, 1990).

Researchers have attempted to develop methodologies that address some of these limitations. For example, the structural analysis of social behavior (SASB) is a measure that was
developed upon a circumplex model of interpersonal behavior. This measure aims to go beyond assessing individual contributions to the therapy process and instead examines the unique system created by the two individuals (Benjamin, Foster, Roberto, & Estroff, 1986). The SASB has also been used to predict how one participant’s behaviors affect the other within subsequent turns in sessions (Benjamin, 2002; Hill, 1990). The coding procedure begins by breaking down session conversations into thought units, which are operationalized as any portion of speech expressing one complete thought. Interpersonal interactions are then coded in terms of focus (other, self, introjection), the degree of interdependence and affiliation, and on nine topics (e.g., approach/avoidance, need for fulfillment, and intimacy). This position is subjected to a global judgment check, which confirms the interpersonal cluster into which the unit has been coded. A strength of the SASB system is the many different levels at which interpersonal process can be analyzed (Samstag et al., 2008). It also permits extremely fine-grained analysis of interpersonal events, and uses small rating units judged by methods requiring relatively low inference and permitting high specificity (Henry et al., 1986). However, it also requires substantial training and resources, at least two coders, and is often applied to only a subset of data or a shortened segment of a session.

Another method developed to identify interpersonal processes occurring in a therapeutic relationship is the Facilitative Interpersonal Skills (FIS) Performance Task (Anderson et al., 2009). This task was designed to measure therapists’ abilities to respond to challenging interpersonal situations in a therapy setting. The performance task consists of eight brief video segments of problematic patient-therapist interactions reenacted with actors. Therapists are prompted to respond to the patient-actors at predefined moments as if they were the therapist in the situation. In the first study (Anderson et al., 2009), two licensed Ph.D. research clinicians
rated each of the eight recorded responses for each therapist. After studying the manual, the two raters (one of which developed the manual) met for two days to discuss and practice ratings with sample responses, before separately rating all study responses. The mean scores for the two raters on each item were then summed to obtain one FIS performance rating for each therapist. This method elucidates detailed and complex information, but requires substantial resources for training and coding the material, which also limits the practical size of the sample. Subsequent research demonstrated that between-trainee differences in FIS scores accounted for significant between-therapist outcome differences in future cases tracked with routine outcome measures (Anderson, McClintock, Himawan, Song, & Patterson, 2016). Importantly, the FIS ratings were assigned based on trainees’ responses to the standardized scenarios; for the reasons detailed above, FIS were not directly measured with the subsequent cases where outcome was monitored.

The Achievement of Therapeutic Objectives Scale (ATOS; McCullough et al., 2003) is a coding measure rated for patient behavior demonstrating the achievement of defined levels of various treatment objectives. These ratings are made from video-recorded segments of sessions, during which every 10 minutes a rating is made according to a 0 to 100 scale, divided into 10 levels. The ATOS comprises seven subscales that assess the degree to which the patient is able to “take in” or absorb the following specific objectives: (a) insight (defense recognition), (b) motivation, (c) experiencing (adaptive) activating affect, (d) experiencing inhibitory affect, (e) new learning (adaptive expression of thoughts, feelings, wishes, or needs), (f) sense of self, and (g) sense of others. Ratings of each subscale are based on the rater’s observation and judgment of the patient’s verbal report or observable bodily reactions (e.g., sighing, crying, tensing, etc.). Similarly, this measure requires extensive training and coding time (Schanche et al., 2010).
A final example of a novel method for studying moment-to-moment interpersonal processes in psychotherapy is joystick coding (Thomas et al., 2014). With this method, an observer uses a computer-connected joystick to make observational ratings of recorded interactions. The interpersonal data appear as nearly continuous flows of a circumplex defined by dimensions of control (dominance/submission) and affiliation (friendly/unfriendly). Results yield a sampling of each person’s interpersonal behavior across a session, as well as information about the temporal dynamics that interrelate clients’ and therapists’ behaviors. Previous literature suggests that interactions are most harmonious (i.e., least anxiety provoking and most stable) when individuals in a dyad behave in a manner that is similar with respect to affiliation but opposite with respect to control—a pattern referred to as complementarity (Sadler & Woody, 2003; Sadler, Ethier, Gunn, Duong, & Woody, 2009; Tracey, 2004). Thus, studying interpersonal complementarity may provide a window into client-therapist relationship patterns that play an important role in treatment. However, this method necessitates training and reliability standards, as well as time and effort coding study data.

To summarize, process-oriented researchers have worked to develop observer-rated measures of interpersonal processes within psychotherapy to better understand the underlying mechanisms of change that lead to successful outcomes for clients. These measures overcome many of the problems that arise from self-report or therapist-report instruments on the therapeutic relationship and/or alliance, including gathering much more detailed, concrete, and objective information. Yet, these measures have their own setbacks. They inherently require a large amount of training, coding time and effort on the part of multiple individuals. They also often necessitate sampling a subset of sessions or segmenting sessions, which surrenders valuable data. Nevertheless, novel tools with the capacity to objectively measure various
interpersonal processes in real time, without a limit on number or hours of sessions one can analyze, hold significant promise for psychotherapy process research.

The Sociometric Badge

Technology innovations have the potential to increase the efficiency of data collection and processing through potentially eliminating the need for human coders. By tracking and recording interpersonal processes in real-time, more detailed, objective data can be collected for complete (rather than segmented) psychotherapy sessions, across a comprehensive series of sessions. Technology can also allow for increased uniformity of measurement, which can be compared more seamlessly across studies of varied interpersonal events. The Human Dynamics Group at the MIT Media Lab has developed a novel electronic device, the Sociometric Badge 03-02 (Sociometric Solutions, 2013). These palm-sized, portable devices are designed to be worn around the neck, and gather data on social signals (e.g., speech pattern, body movement) in real-time. The “badge” design is intended to support its use in ecologically valid circumstances. They are capable of tracking face-to-face interaction, conversational time, physical proximity to other people, and physical activity level by measuring vocal features, body motion, and relative location. When multiple badge wearers are in proximity, built in infrared sensors “ping” and sync the badges, which allows for the study of multivariate, interdependent dyadic data.

However, Sociometric Badges (SBs) have never been utilized in psychotherapy research or within psychotherapy sessions.

The badges were developed based on research examining “Honest Signals” (Pentland & Heibeck, 2010). Pentland defines honest signals as indicators that humans unknowingly exchange that convey intents, emotional states, and personal characteristics. Several representations of honest signals can be derived from observing the give-and-take of
conversational turn-taking and gesturing, and carefully measuring the timing, energy, and variability of the interaction. For example, *influence* in social interaction, indicating dominance versus submission, is measured by the extent to which one person causes the other person’s pattern of speaking to match their own pattern. Research has demonstrated that strength of influence in a conversation serves as an honest signal of attention, and that people typically use this signal to assess attitudes and interest level of others (Pentland & Heibeck, 2010). Another process known as *mimicry* or the “chameleon effect” refers to the reflexive copying of one person by another during a conversation, such that one’s behavior passively and unintentionally changes to match that of others in one’s social environment, resulting in unconscious back and forth trading of smiles, interjections, and head nodding (Chartrand, Maddux, & Lakin, 2005; Chartrand & Bargh, 1999). This construct is particularly relevant for psychotherapy because of its association with conversational partners’ self-report of trust and liking of the other. Negotiations where participants unconsciously mimic one another tend to be smoother and more successful (Chartrand et al., 2005). In addition, empathic people are more likely to mimic their conversation partners; thus, mimicry is often described as an unconscious signal of empathy. Empathy has been associated with both the quality of the working alliance and treatment outcome (Ackerman & Hilsenroth, 2003; Feller & Cottone, 2003). Within psychotherapy, this information could serve as helpful feedback to clinicians in understanding how their patients perceive their behaviors and the relationship.

Pentland and Heibeck (2010) further describe that people in real-life situations employ combinations of honest signals. These combinations cluster in a manner that infers the social role a person has assumed—specifically the attitudes, intentions, and goals that exist in the relationship. For instance, the social role of exploring is exhibited by displaying a combination of
honest signals of interest and an openness to influence, while active listening encompasses a display of attentive interest and openness to new ideas. These roles have then been used to accurately predict the outcomes of interactions (e.g., successful negotiation of a problem; Pentland et al., 2010). Honest signals are considered somewhat separate from other modes of expression (e.g., language, emotional display, culturally defined gestures) in terms of their time course, being much slower than words or sentences.

Types of data captured by the SBs include speech and movement energy, speech frequency, and group interaction patterns, among other real-time emergent collaboration patterns on a large scale. Its current capabilities include measuring human movement using a single 3-axis accelerometer to detect activities such as walking, sitting, nodding, and hand movement; extracting speech features in real time to measure nonlinguistic social signals such as enthusiasm, interest level and persuasiveness; sending and receiving information over 2.4 GHz radio to and from different users and base stations; performing indoor user localization by measuring received signal strength from fixed base stations; capturing face-to-face interaction time using an infrared transmission (IR) sensor, so that when badge wearers have a direct line of sight to each other, an IR signal will be received; capturing proximity data by using Bluetooth and the radio transceiver; and communicating with Bluetooth enabled mobile devices to provide feedback to the user. Analyzing the data with the Sociometric Solutions software provides output that reveals body movement analysis (activity, consistency, mirroring, influence, and rate), posture analysis (activity, rate, and mirroring), speech analysis (activity, consistency, mirroring, influence, frequency, speech profile—percentage of speaking/listening/silent/overlap, and dominance), turn taking analysis (including number of turns, speaking segments, pauses,
successful/unsuccessful interruptions, and average speaking segment and pause length), and social network analysis (face-to-face detection, network cohesion, proximity matrix).

Tracking and analyzing these signals in a psychotherapy setting may elucidate useful information pertinent to the therapeutic alliance and interpersonal dynamics within a given dyad, and provide data that may help guide facilitative clinician behavior. SBs utilize a technology that enables automatic measurement of collective patterns over time and across physical boundaries at a scale that has previously been impossible in group or process research (Kim, McFee, Olguin, Waber, & Pentland, 2012). They also represent a potentially great advantage over direct observation by humans (Olguin & Pentland, 2007). Deploying pervasive cameras is also extremely expensive and constrains range of measurement to a particular place. The ability to automatically capture visible characteristics of humans and the underlying psychological processes that occur during social interactions in hundreds of people, at the same time, with a single unobtrusive tool, is less feasible with other methods.

Research to date utilizing the SBs has been conducted in organizational settings to attempt to enhance organizational behavior (e.g., it is designed to be worn around one's neck like a typical company ID badge). Researchers use the badges to measure individual and collective patterns of behavior, predict human behavior from the social signals collected, identify social affinity among individuals working in the same team, and enhance social interactions by providing feedback to the users of the system. Kim, Chang, and Pentland (2007) discuss the potential implications of providing a solution to questions such as how individuals can get regular feedback on their communication habits, and how leaders can enhance the effectiveness of their teams. In turn, they propose using the Sociometric badges in meetings to help participants better understand the flow of the meeting and improve participation. For example, as
the badges detect the turn-taking pattern of the conversation, they can alert participants if they are dominating a conversation, while giving other people reminders to participate more in the conversation. The data can also inform people on their speaking style in meetings by detecting individual activities such as gestures and posture, as well as social interactions such as body movement mimicry or rhythmic patterns. If one’s speech always leads to boredom or an emotionally negative response (e.g., fidgeting behavior), they may receive a recommendation for a change in their manner of speech.

In the closest study of relevance to health care applications to date, Olguin, Gloor, and Pentland (2009) studied a group of 67 nurses working in the Post Anesthesia Care Unit of a Boston area hospital. Using the data collected with the badges worn on the unit, they identified different individual personality traits. For example, they found that higher daily percentage of face-to-face time and more variation in this percentage day-to-day indicated higher neuroticism in nurses. The characteristic of agreeability was exhibited by less variation in average speech volume modulation and less variation in daily percentage of time in close proximity to a bed. This study confirmed the feasibility of identifying individual personality traits from sensor data. In addition, they were able to estimate the overall group performance by aggregating the daily sensor features across subjects.

**Current Study**

When feasible, the traditional approach to human coding involves the rating of session recordings. This is typically done through audio and video recording devices. A common approach is to place a digital recorder next to or in between therapy participants. Outside of a research context, this approach is commonly employed in therapist-training settings. Patient consent for recording may even be a requirement to be seen for therapy in such settings. This
approach is not without controversy. Briggie, Hilsenroth, Conway, Muran, and Jackson (2016) found that patient comfort with recording was not statistically related to treatment refusal, duration, or outcome in a sample of nearly 400 outpatients treated in a university-based community mental health setting. Approximately 50% of patients expressed no or slight concerns about recording, and higher levels of interpersonal sensitivity and paranoia were inversely related to comfort.

For the reasons outlined above, Sociometric badges hold promise for utilization in psychotherapy process research. However, researchers and clinicians alike may find it important to understand if the badges will cause distraction in session or otherwise negatively impact the relationship quality before implementation in real-world treatment settings. Therefore, the current study piloted the use of the badges in an undergraduate sample to preliminarily assess the acceptability and feasibility of these assessment devices in comparison with traditional audio recording equipment in a context that was similar to individual psychotherapy. Concretely, participants were assigned to dyads on a rolling basis and instructed to engage in a series of speaking tasks meant to represent a psychotherapeutic dialogue. Each dyad was randomly assigned to a condition where both participants wore a badge or a condition with a digital recording device in the room and no badges. After each conversation task, both participants were asked to rate their experience of their partner; at the end of the tasks, both participants were asked about their impressions of the recording device that was used in their condition.

The study aimed to examine differences in the perceived quality of the interaction depending on the nature of the assessment device, as well as dyad satisfaction with the specified device. It was hypothesized that dyads would not differ on perceived quality of interaction or device satisfaction by condition. The study also explored whether global interpersonal distress,
gender match and race match of the participants, and conversation topic valence (neutral, positive, or negative) influenced participant liking of the conversation partner.

**Method**

**Participants**

Participants were undergraduate students (n = 306) at a large public University in New York state who participated for research credit in an introductory psychology course. They were recruited online through a university research pool. Eligible participants were 18 years of age or older, and were paired into dyads (n = 153 dyads) for the purpose of this study. Each dyad was randomized into the Sociometric Badge condition (n = 75 dyads) or the traditional audio-recorder condition (n = 78 dyads). The mean participant age was 19.16 years (SD = 1.44, range = 18 to 29 years). The sample was mostly female (50.3%) and Caucasian (47.1%), followed by African-American (17.3%), Asian (15.0%), and Hispanic (13.7%). The most common academic standing was freshmen (48.9%), followed by sophomores (31.4%), juniors (13.7%), and seniors (6.2%). The Sociometric badge condition included 48.0% of dyads matched on gender and 32.0% matched on race, while the recorder condition included 47.4% of dyads matched on gender and 30.8% matched on race. No significant differences in demographic characteristics were found between participants in the badge versus recorder conditions.

**Measures and Materials**

**Self-Report Measures**

**Inventory of Interpersonal Problems** (IIP-64; Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988). This 64-item self-report instrument measures interpersonal style and the degree of distress arising from interpersonal problems. Responses for each item range from 0, *Not at all*, to 4, *Extremely*. Example items include, “It is hard for me to: Trust other people,” “It
is hard for me to: Really care about another person’s problems,” and “The following things are things you do too much: I feel embarrassed in front of other people too much.” A total score can be derived as an index of overall interpersonal distress, which was used in the present study. The 64-item IIP has demonstrated strong reliability and validity, and is sensitive to clinical change (Horowitz et al., 1988).

**Relationship Quality.** This 7-item self-report measure was developed for the purpose of this study by study authors. Participants were asked to rate items on a scale from 1, *Not at all*, to 5, *Very much*, regarding their momentary feelings about their partner in the dyad. Items included questions such as, “How much do you like your partner?”; How much do you think your partner likes you?”; “Can you see yourself becoming friends with this person?”; and “How enjoyable did you find this most recent interaction?” However, one question, “How upsetting did you find this most recent interaction?”, was eliminated from analyses due to its inadequate psychometric performance, possibly due to its reverse coding. This measure was administered at three time points across the interaction (once after each dyadic conversation type described below). The 6 remaining items were summed to create a total relationship quality score (ranging from 6 to 30) for each participant and for each conversation (neutral, positive, and negative).

**Device Satisfaction.** This 4-item self-report measure was also developed by study authors for the purpose of this study. Participants were asked to rate their general satisfaction with the assigned recording device (Sociometric badge vs. recorder) on a scale of 1, *Not at all*, to 5, *Very much*. Specific items asked, “How comfortable were you with the recording device?”; “How much did the recording device interfere with your ability to interact normally?”; “How aware were you of the recording device during the conversations?”; and “How interesting did you view the recording device?” This questionnaire was administered once, after all
conversations had been completed. Each question was analyzed separately (a total score was not created for this measure).

**Sociometric Badge**

The *Sociometric Badge 03-02* (Sociometric Solutions, 2013) is a palm-sized, portable device designed to be worn around the neck. These badges gather data on social signals (e.g., speech pattern, body movement) in real-time. The “badge” design allows for use in ecologically valid circumstances (see the picture in Figure 1), such as an organizational setting or between partners during an interpersonal interaction. Each participant in a dyad randomized to the badge condition wore a Sociometric Badge around their neck during the contrived conversations.

Although the present paper did not focus on the findings from the badge data collected, future research will examine participant body movement analysis, posture analysis, speech analysis, and turn taking analysis in relationship to the perceived quality of the relationship and interpersonal style and distress in this subgroup.

**Procedures**

All study procedures and materials were approved by a University institutional review board (IRB). Participants enrolled for the study online, and were provided a specific time slot to arrive at the lab. Once both partners of the dyad arrived, a research assistant greeted them, provided the informed consent form, explained the study procedures, and answered any questions. The study was conducted in two separate labs, each with an identical room set up with two chairs facing one another and a small table in between. Alternating dyads were randomized into the Sociometric Badge condition (i.e., each participant wore a badge around their neck) versus the traditional audio-recorder condition (i.e., the audio recorder was placed on a table between the participants). After signing the consent form, participants completed the initial self-
report measures. In addition, they were asked to write down a description of both a negative and positive emotional memory they could discuss with their study partner for three minutes each.

Research assistants used a standardized script to explain study procedures which were as follows: participants will engage in one conversation about a neutral topic for three minutes, specifically a conversation about standardized tests (e.g., Scholastic Assessment Test). Next, each participant had three minutes to discuss his/her positive memory, and then his/her negative memory while the other partner listened and engaged the speaker about this memory. For example, after the neutral conversation, partner 1 had three minutes to talk about his/her positive memory while partner 2 listened and collaboratively discussed this memory. Then, partner 2 discussed his/her positive memory while partner 1 listened and collaboratively discussed this memory. This procedure was then repeated for the negative emotional memory. The research assistant flipped a coin to determine if the positive or negative memory conversation would be first or second to control for an ordering effect. Each conversation was timed by the research assistant for accuracy. At each time point after the neutral, positive, and negative conversations ended, the participants completed the relationship quality questionnaire (a total of three times). After all conversations were concluded, participants completed the device satisfaction questionnaire. Participants were granted research credit for their participation.

Data Analytic Plan

In order to determine the most appropriate inferential test approach for the study-specific measure of relationship quality, a principal components analysis (PCA) of the relationship quality instrument items was conducted for each administration of the measure (following the neutral, positive, and negative conversations), in order to identify the number of components empirically driving the participant responses. A PCA was the chosen method of extraction
because the goal was to empirically summarize the data (Tabachnick & Fidell, 2013). There was no underlying theoretical factor structure; rather, we were primarily interested in reducing these variables down to a smaller number of components, using maximum variance from the data.

Hierarchical linear modeling (HLM) was utilized to test differences in relationship quality ratings between the two conditions. This mixed model included random dyad intercepts because data were nested, such that individuals were nested within dyads. Other variables that were speculated to impact an individual’s liking of another were controlled for, including global interpersonal distress (IIP-64 total score), dyad match/mismatch of gender and of race, and conversation topic. In addition, a multivariate analysis of variance (MANOVA) was conducted to test differences in device satisfaction items between conditions. A MANOVA was chosen as it potentially protects against inflated Type I error and can reduce Type 2 error when dependent variables are moderately correlated (Tabachnick & Fidell, 2013). The nature of the device questions (limited number of items and content) suggested that it would be most appropriate to treat each item as a separate DV.

Results

Preliminary Analyses

A small degree of random item-level missingness on the IIP-64 measure was observed. The missing value analysis indicated that no given item had more than 2.5% missing values. Thus, these missing data points were imputed using a maximum likelihood procedure, estimation maximization (Allison, 2012). Participants missing an entire instrument were eliminated from the relevant analyses. Next, a descriptive analysis of the data was conducted; normality assumptions were met, as skewness and kurtosis for all variables of interest were within normal limits. One case was found to be a univariate outlier on the IIP-64 total interpersonal distress
score, and consequently this data point was winsorized. No multivariate outliers were found according to the Mahalanobis distances. See Table 1 for descriptive statistics of key variables.

Three separate PCAs were performed on the 6 items of the relationship quality questionnaire given at each time point (after the neutral, positive, and negative conversations) with an oblique rotation. For both the positive and negative conversations, one component was extracted from the 6 items. The determinants were greater than 0, indicating factorability of the data. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was significant for both PCAs. Similarly, for both anti-image correlations, the on-diagonal values were large (< .75), while the off-diagonal values were small (> .46), suggesting acceptable factorability. Extracted communalities were all above .4 (except for the “How helpful did you find this most recent interaction?” item after the positive conversation, which was .349). In addition, both PCAs evidenced one component with an eigenvalue greater than one, consistent with the scree plot. A parallel analysis additionally supported a one factor model. All items in both PCAs loaded onto one component at a value of .59 or higher.

The PCA on the relationship quality items for the neutral conversation originally loaded onto two components. Although all other factorability indices mentioned above were adequate, the pattern matrix revealed a complex structure, in which some items loaded onto both components. Therefore, another PCA was performed for which items were forced into one component. All factorability indices remained adequate, and items loaded well on to the single component (all values above .55). The parallel analysis was also consistent with a one component model. Therefore, in determining the appropriate score to use as the dependent variable, we decided to sum all relationship items from each conversation for one overall total relationship quality score.
Perceived Relationship Quality

HLM was used to test differences between the Sociometric badge and audio recorder conditions on the relationship quality instrument, controlling for interpersonal distress, gender match and race match of the dyadic partners, and conversation type (see Table 2 for results). Eleven dyads reported knowing one another prior to participation in the study, which may have impacted their relationship quality ratings. However, excluding these dyads from analyses did not change the nature of the results, and these dyads were therefore retained in the analysis. As expected, relationship quality did not differ between conditions above and beyond the control variables ($b = .43, p = .38$). Relationship quality was also not impacted by a match or mismatch of race of the two dyadic partners ($b = -.40, p = .45$). However, gender match of the dyadic partners ($b = -1.37, p < .01$), conversation topic ($b = -.27, p < .001$), and interpersonal distress ($b = -.53, p < .05$) were all significant predictors of relationship quality. Specifically, participants in gender mismatched dyads and individuals endorsing greater interpersonal distress exhibited significantly lower average relationship quality ratings. Furthermore, participants reported lower average relationship quality following the neutral conversation than following the positive or negative conversations. The positive and negative conversations did not differ from one another.

Device Attitudes

A MANOVA was conducted to test mean differences between conditions on the four device attitude/satisfaction questionnaire items (see Table 3 for results). The pooled determinant was greater than zero ($D = 1.95$), indicating that singularity was not present. Box’s M was not statistically significant ($Box’s M = 10.48, p = .412$), which suggested adequate homogeneity of variance-covariance. The within residual correlations were small, all of which were equal to or below .47. As such, all assumptions were met. The overall multivariate test was significant,
Wilk’s $\Lambda = .936$, $F(4, 298) = 5.09$, $p < .001$, revealing a difference between conditions on one or more of the device satisfaction items. The only statistically significant univariate test was for device question three, $F(1, 298) = 10.03$, $p < .01$, which asked, “How aware were you of the recording device during the conversations?” Interestingly, participants in the standard audio recorder condition ($M = 2.77$) rated that they were more aware of the device than participants in the Sociometric badge condition ($M = 2.32$).

**Discussion**

The present study aimed to assess the acceptability and feasibility of a novel technological tool, the Sociometric Badge, as a device for measuring a wide range of interpersonal process variables during a simulated psychotherapeutic dialogue between undergraduate students. Before implementing this tool in an authentic psychotherapy setting, the authors sought to test if the perceived quality of the relationship or interaction between analogue therapy participants would be compromised by the Sociometric badge, as compared to traditional audio recording equipment already commonly used in the field. We first tested if the dyads that wore these badges rated perceived relationship quality with their partners differently than dyads being recorded with a traditional audio recorder in the room. As hypothesized, no differences in relationship quality were found between conditions, after controlling for other potential predictors. These findings suggest that in a simulated setting these badges do not differentially (or negatively) impact perceptions of the interpersonal interaction compared to a standard audio recording device.

Some of the variables controlled for in this analysis did significantly predict one’s perceived quality of the relationship and liking of the other participant. Dyads that differed in gender demonstrated lower average relationship quality than those comprised of the same
gender. It is possible that in the instance of sharing a personal positive and negative emotional memory with an unfamiliar other, individuals felt greater comfort with a partner of the same gender, which contributed to higher ratings of liking and closeness. Yet, previous research examining the association between gender match and alliance in psychotherapy is mixed, with some studies finding that gender match leads to a stronger alliance (Johnson & Caldwell, 2011; Wintersteen, Mensinger, & Diamond, 2005), others finding that a match was more important to the alliance for female clients (Bhati, 2014), while other results have demonstrated a more distressing disclosure experience for female patients working with a gender-matched therapist (Pattee & Farber, 2008).

Moreover, in the present study relationship quality did not differ significantly based on a match or mismatch of identified race/ethnicity. Studies of patient-therapist race match on patient retention, alliance, and outcome in actual psychotherapy have been inconclusive and lack a foundation of rigorous research designs (Karlsson, 2005). Of the few studies that are available, most suggest that matching on ethnicity does not affect the outcome of therapy (e.g., Sterling, Gottheil, Weinstein, & Serota, 2002). Wintersteen and colleagues (2005) found that matching on race predicted greater patient retention, but not patient-rated alliance. Therapists in this study, however, believed that being matched with a patient of the same race would lead the patient to feel a stronger alliance with them. Overall, the field lacks strong empirical support for matching based on ethnicity in psychotherapy. More research examining the importance of a gender-matched and/or race/ethnicity-matched therapist for a particular client is needed to more clearly understand these relationships.

In addition, higher self-reported interpersonal distress predicted lower relationship quality ratings. Literature on the impact of interpersonal difficulties on the alliance reports similar
results. For example, research with different therapeutic orientations and varying clinical populations has consistently found that interpersonal distress, and more specifically hostile-dominant interpersonal problems, significantly predicts poor alliance in treatment (Constantino & Smith-Hansen, 2008; Dinger, Strack, Sachsse, & Schauenburg, 2009; Gibbons et al., 2003; Muran, Segal, Samstag, & Crawford, 1994). In this study, perhaps individuals with greater interpersonal distress experienced anxiety during the personal conversations and struggled with thoughts of negative self-evaluation, or alternatively were untrusting of their partner. These thoughts may have led individuals to believe that their partner does not like them or cannot be trusted, ultimately leading to lower ratings of the relationship.

The conversation topic type was also a significant predictor of relationship quality, such that participants rated the relationship quality lower on average following the neutral conversation than following the positive or negative emotional memory conversations. Given that the neutral conversation was three minutes in duration during which participants discussed their views on standardized testing, it may not be surprising that a weaker relationship quality was found. This was to some degree expected by the researchers as the neutral conversation was intended to be a control or practice/baseline phase before the more personal and potentially evocative conversations took place. Once one is introduced to more personal memories of another individual’s life, one may naturally feel a greater sense of closeness and liking. Research findings support the relationship between self-disclosure and feelings of relational closeness and satisfaction (Sprecher & Hendrick, 2004).

The authors additionally sought to examine if dyads in the badge condition rated their opinions of these devices differently than dyads in the traditional audio recorder condition. Consistent with hypotheses, three of the four device satisfaction items related to comfort with the
device, interference of the device in one’s ability to interact normally, and how interesting one viewed the device, did not differ between study conditions. These results provide further evidence that using the Sociometric badges in a psychotherapy context is unlikely to impact the experience of the participants above and beyond their typical experience with an audio recorder in the room. However, ratings of the item, “How aware were you of the recording device during the conversations?” did differ between conditions, such that those in the traditional audio recorder condition endorsed being more aware of the device during the interaction than those in the badge condition. Although this difference was somewhat unexpected, it appears to further support the use of the badge in such research. Participants may have noticed more clearly the tape recorder sitting on the table than the smaller white badges around their necks. This is in line with the intent of the badge developers, which was to mimic the type of ID badge that a company employee might wear, but largely forget about wearing throughout most of their daily activities.

Overall, using a Sociometric Badge as an assessment tool does not seem to interfere with perceived relationship quality, or cause significantly more distraction or interference than a traditional recording device in a simulated psychotherapeutic dialogue. Moreover, these badges have the technological power to innovate and expedite within-session process research for more efficient research than possible with human coders. Process research utilizing observational methodologies requires significant time in training coders to reliability, as well as the time and effort required to code and come to consensus on study data. These requirements often lead to the use of session and/or segment sampling, which may contribute to the loss of valuable data. Human error and implicit rater biases are always present as well. Conversely, these badges offer objective and automatic measurements of time series data, over the entirety of a session (as well as across multiple sessions in terms of battery life and memory capacity). Their portability
allows them to be used in ecologically valid settings, such as the therapy room, for therapy dyads or groups with any number of members. The hardware integrates with software by connecting the badge to a computer and uploading the data to a spreadsheet or a statistical program of one’s choosing. This integration offers easy accessibility to summary statistics on variables of interest.

Limitations

The findings from the present study must be considered in light of the study’s limitations. First, this study was not conducted in an authentic psychotherapy setting, rendering the generalizability of the results to applied clinical practice speculative. Nevertheless, it was important – both methodologically and ethically – to pilot this research in a simulated setting to determine its acceptability in a similar environment to gain understanding about its impact on interpersonal interaction and relationship and rapport building. These findings are also limited by the time constraint of the study, which consisted of a total of 15 minutes of conversation time. This restricted amount of time may not be sufficient in creating a reliable and valid relational bond and determining one’s liking and feeling of closeness to another individual.

Furthermore, the sample consisted of an undergraduate population enrolled in an introductory psychology course, additionally restricting the generalizability of the results. The younger student population may have a skewed viewpoint on the use of new technologies, as well possess a greater familiarity and comfort with such technologies than older or non-student populations (Czaja et al., 2006). Lastly, the authors created the relationship quality and device satisfaction questionnaires for the purpose of this study, and thus lack comprehensive research on their psychometric properties, including the adequacy of their reliability and validity.

Conclusion and Future Directions
Despite the limitations of the current study, the acceptability and feasibility of the use of the Sociometric Badge in a psychotherapeutic setting is promising. The use of these devices does not seem to impact the quality of the relationship or interfere in wearers’ abilities to interact normally compared to the use of a traditional audio recorder. In the future, the researchers would like to examine the Sociometric badge data from the relevant subsample, to observe how the measurements correlate with self-reported interpersonal distress, relationship quality ratings, and device satisfaction ratings. An examination of the descriptive statistics from the badges on variables such as turn taking, number of speaking segments, silence, volume modulation, body movement, posture analysis, and mirroring of these variables between the two dyadic partners is also of interest. Future research should pilot these badges in a true psychotherapy setting to further assess their acceptability and feasibility with a clinical population. This research would ultimately seek to better understand if the honest signals measured by this tool correlate with or predict alliance ratings and/or psychotherapy outcomes.
References


interpersonal psychology: Theory, research, assessment, and therapeutic interventions
(pp. 509-518). New York: Wiley.


Constantino, M. J., Boswell, J. F., Coyne, A. E., Kraus, D. R., & Castonguay, L. G. (2017). Who works for whom and why? Integrating therapist effects analysis into psychotherapy outcome and process research. In L. G. Castonguay & C. E. Hill (Eds.), How and why are
some therapists better than others? Understanding therapist effects (pp. 55-68).


### Descriptive Statistics of Key Variables by Condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Item</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recorder</strong></td>
<td>IIP Total</td>
<td>154</td>
<td>77.37</td>
<td>36.23</td>
</tr>
<tr>
<td></td>
<td>Relationship Quality: Neutral</td>
<td>156</td>
<td>20.92</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>Relationship Quality: Positive</td>
<td>154</td>
<td>23.88</td>
<td>3.83</td>
</tr>
<tr>
<td></td>
<td>Relationship Quality: Negative</td>
<td>154</td>
<td>23.71</td>
<td>3.74</td>
</tr>
<tr>
<td></td>
<td>Comfort with device</td>
<td>154</td>
<td>3.97</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>Interference of device</td>
<td>154</td>
<td>2.05</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>Awareness of device</td>
<td>154</td>
<td>2.77</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>Device interesting</td>
<td>154</td>
<td>2.61</td>
<td>1.21</td>
</tr>
<tr>
<td><strong>Badge</strong></td>
<td>IIP Total</td>
<td>149</td>
<td>77.80</td>
<td>35.87</td>
</tr>
<tr>
<td></td>
<td>Relationship Quality: Neutral</td>
<td>150</td>
<td>20.65</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td>Relationship Quality: Positive</td>
<td>150</td>
<td>23.16</td>
<td>4.01</td>
</tr>
<tr>
<td></td>
<td>Relationship Quality: Negative</td>
<td>150</td>
<td>23.16</td>
<td>4.20</td>
</tr>
<tr>
<td></td>
<td>Comfort with device</td>
<td>149</td>
<td>3.88</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>Interference of device</td>
<td>149</td>
<td>1.88</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>Awareness of device</td>
<td>149</td>
<td>2.32</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>Device interesting</td>
<td>149</td>
<td>2.87</td>
<td>1.18</td>
</tr>
</tbody>
</table>

*Notes. IIP = Inventory of Interpersonal Problems; Neutral, Positive, Negative refers to conversation type.*
Table 2.

*Results of Hierarchical Linear Model of Predictors of Relationship Quality*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>SE</th>
<th>df</th>
<th>t</th>
<th>p</th>
<th>CI: Lower</th>
<th>CI: Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>24.20</td>
<td>0.56</td>
<td>156.92</td>
<td>43.52</td>
<td>.000</td>
<td>23.10</td>
<td>25.30</td>
</tr>
<tr>
<td>Gender match</td>
<td>-1.37</td>
<td>0.49</td>
<td>148.83</td>
<td>-2.78</td>
<td>.006</td>
<td>-2.34</td>
<td>-3.94</td>
</tr>
<tr>
<td>Race match</td>
<td>-0.40</td>
<td>0.53</td>
<td>147.80</td>
<td>-0.76</td>
<td>.451</td>
<td>-1.44</td>
<td>0.64</td>
</tr>
<tr>
<td>Trial</td>
<td>-2.67</td>
<td>0.17</td>
<td>594.61</td>
<td>-15.91</td>
<td>.000</td>
<td>-3.01</td>
<td>-2.34</td>
</tr>
<tr>
<td>Condition</td>
<td>0.43</td>
<td>0.49</td>
<td>148.47</td>
<td>0.88</td>
<td>.380</td>
<td>-0.54</td>
<td>1.40</td>
</tr>
<tr>
<td>IIP</td>
<td>-0.53</td>
<td>0.25</td>
<td>238.83</td>
<td>-2.13</td>
<td>.034</td>
<td>-1.01</td>
<td>-0.04</td>
</tr>
<tr>
<td>Condition*IIP</td>
<td>0.66</td>
<td>0.36</td>
<td>256.15</td>
<td>1.86</td>
<td>.065</td>
<td>-0.04</td>
<td>1.36</td>
</tr>
</tbody>
</table>

*Note.* IIP = Inventory of Interpersonal Problems general score. Condition = badge vs. standard recording device
Table 3.

*Multivariate Analysis of Variance of Device Satisfaction*

<table>
<thead>
<tr>
<th>Wilk’s Λ</th>
<th>Value</th>
<th>Exact F</th>
<th>df</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.936</td>
<td>5.087</td>
<td>4</td>
<td>.001</td>
<td>.064</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Univariate</th>
<th>SS</th>
<th>Error SS</th>
<th>F</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Device Comfort</td>
<td>0.591</td>
<td>342.660</td>
<td>0.519</td>
<td>.472</td>
<td>.001</td>
</tr>
<tr>
<td>Device Interference</td>
<td>2.093</td>
<td>388.507</td>
<td>1.622</td>
<td>.204</td>
<td>.005</td>
</tr>
<tr>
<td>Device Awareness</td>
<td>15.836</td>
<td>475.220</td>
<td>10.030</td>
<td>.002</td>
<td>.032</td>
</tr>
<tr>
<td>Device Interest</td>
<td>4.939</td>
<td>427.939</td>
<td>3.474</td>
<td>.063</td>
<td>.011</td>
</tr>
</tbody>
</table>
Figure 1.

Sociometric Badge