Do You Even Need Sensors?: Synthetic Biomusic as an Empathic Technology

Daway Chou-Ren
Microsoft Research
San Francisco, CA, USA
dawaychouren@gmail.com

Mike Winters
Microsoft Research
Redmond, WA, USA
mikewinters@microsoft.com

Javier Hernandez
Microsoft Research
Cambridge, MA, USA
javierh@microsoft.com

Daniel McDuff
Microsoft Research
Seattle, WA, USA
djmcduff@media.mit.edu

Jina Suh
Microsoft Research
Redmond, WA, USA
jinsuh@microsoft.com

Vanessa Rodriguez
Microsoft Research
Redmond, WA, USA
vanessa.rodriguez@microsoft.com

Gonzalo Ramos
Microsoft Research
Redmond, WA, USA
goramos@microsoft.com

Mary Czerwinski
Microsoft Research
Redmond, WA, USA
marycz@microsoft.com

***Abstract*—Previous research suggests that biomusic, a type of biosignal sharing, is effective at promoting empathy and closeness among individuals. However, it is unclear whether these effects are due to the information it encodes or other emotional aspects of its resulting music. To explore this question, we developed a Generative Adversarial Network (GAN) to create synthetic biomusic that approximates real biomusic, and employed deception to evaluate its effects on 24 pairs of participants engaged in real-time emotional disclosure. Users reported that both real and synthetic biomusic provided the same amount of information about their conversational partner as observing body language, facial expressions, or vocal tone. Further, both conditions increased users' ratings of closeness and empathy with each other compared to listening to no music. However, we found no statistically significant differences between the two biomusic conditions across any of our metrics. We discuss the implications of these results for the design of future biomusic systems.**

***Index Terms*—biomusic, biosignal sharing, empathy, closeness, physiology, heart rate, breathing rate, electrodermal activity**

I. Introduction

How to help people connect has become a question of profound importance in recent years, as there is evidence that people have been growing increasingly isolated, disconnected, and lonely even before the COVID-19 pandemic. The number of Americans reporting that they have no close friends is 12% in 2022, up from 3% in 1990. 17% report they are not satisfied with the number of friends they have [1]. One-third of people in industrialised countries "perceive [themselves] to be socially isolated even when among other people" [2].

Real-time biosignal sharing, or exposing physiological information such as heart rate (HR), breathing rate (BR), electrodermal activity (EDA), or other signals to another person, has been found to have promise for increasing the intimacy and empathy that users feel with one another [3]–[7]. While some studies have found these effects with real physiological signal sharing [4], [5], [7], others have found them with synthetically generated sounds that are not linked to real biosignals [3], [6]. This raises the question of whether the connective results of biosignal sharing are due to the specific, real information that's communicated about a subject's physiology or whether these results are due to other factors, such as the *belief* that information is present, users' engagement with the environment created by the biosignal sharing interface, or further confounders (e.g., in the case of biomusic sharing, the intrinsic emotional qualities of the music).

To investigate this, and to further understand and quantify the connective effects of biosignal sharing, we design and develop a biomusic system for sharing sonified biosignals. This system includes the novel use of a time series generative adversarial network (TimeGAN) [8], which allows us to generate a synthetic biomusic that is statistically and auditorily indistinguishable from real biomusic generated from a user's real-time physiology. The synthetic biomusic differs in that it does **not** relate to a user's actual physiology or environment in any way. We compare both biomusics to each other and to a baseline of playing no music at all.

The main research question we try to address is:

- **RQ1** How does the real vs synthetic nature of biomusic (containing real vs synthetic information) change its effects on empathy and closeness?

We also address two other questions:

- **RQ2** How much information do users gain from real and synthetic biomusic compared to traditional information channels such as body language, facial expressions, and tone of voice?
- **RQ3** How do people relate to and use both real and synthetic biomusic?

In this study, we focus on empathy as "the capacities to resonate with another person's emotions, understand his/her thoughts and feelings" and "respond with the appropriate prosocial and helpful behavior" [9].

We run an in-person within-subjects lab study with 24 dyads. Dyads take turns listening to and sharing sad memories with each other while hearing the other's real biomusic, synthetic biomusic, or a no music baseline. We make three key contributions: 1) we investigate the effect of synthetic vs real biomusic sharing, 2) we quantify the effects of both biomusics

979-8-3503-2743-4/23/$31.00 ©2023 IEEE

on measures of closeness and empathy compared to a no music baseline condition, and 3) we investigate how users relate to biomusic and the information they gain from it.

Our study is novel because it explicitly compares synthetically generated biosignal and real biosignal sharing, because we use a real-time bidirectional biomusic system in contrast to previous unidirectional biomusic systems, because we study dyads engaged in an in-person, emotionally intense task while in a laboratory-controlled setting.

II. Related Work

A. Social Effects of Biosignal Sharing

Previous research indicates that exposing a person to physiological signals that are attributed to another person increases intimacy, empathy, and other connective metrics (e.g., Inclusion of Other in Self scale (IOS) [10]). Janssen et al. demonstrate that augmenting interactions with VR avatars by exposing auditory heartbeats increases perceptions of intimacy [3]. Winters et al. show that hearing the heartbeats of another person causes a change in emotional perspective and increases emotional convergence [6]. Other studies, whether researching sharing in VR, chats, or through apps, also find that heartbeat sharing increases empathy or contributes toward connection and intimacy [4], [5], [11]–[13]. While the majority of research has focused on heartbeat sharing, other studies find links between sharing skin conductance [7], [14] and breath rate [14], [15] with empathy and connectedness.

B. Biomusic

Grond et al. define biomusic as "affective technology that communicates emotional states by translating physiological signals into auditory output" [16]. By choosing congruent musical cues to represent physiological signals (e.g., tempo for heart rate), biomusic creates a data sonification that uses the emotional associations of music to enhance its empathic effects [6].

Biomusic has been found to be effective at helping users be more aware of another person's emotional states, for improving their accuracy in classifying them, for increasing feelings of interpersonal connection, for enabling new forms of connection, and for increasing senses of reciprocity and co-presence [14], [16], [17].

The biomusic in these studies is unidirectional, meaning only one user in a dyad has their physiological signals encoded and heard by the other user. We are unaware of research into bidirectional biomusic systems.

C. Synthetic Biosignal Sharing

In most of the prior research, the biosignals shared have been "real," which we define as being mapped directly from actual real-time physiological data. Two exceptions are Janssen et al. and Winters et al., whose findings use pre-generated heartbeats with non-human virtual avatars [3], [6]. However, these studies had simple heartbeats which would not appear out of sync with the behaviors of their virtual avatars. Our synthetic biomusic, while designed to appear statistically and auditorily similar to real biomusic, is not derived from a participant's physiology and is thus not in sync with it though an in sync pattern may still appear by chance. To our knowledge, there have been no studies comparing the effects of "out of sync" synthetic biosignal sharing with the effects of real biosignal sharing. None of the previously cited biomusic literature involved synthetic biosignals.

D. Physiological Signals as Proxies for Affective States

HR and EDA have a well-established history as being among the most relied on proxies for affective states as they are strongly connected to the functioning of the autonomic nervous system [9], [18]–[21]. BR is less well-established but when speakers are sad, their speech slows which changes their breathing rate [18]. These three signals and their variations are also all associated with social interaction dynamics [22]–[24].

We note that these signals cannot represent a user's exact affective state. For example, a high heart rate can indicate anger, fear, or sadness, and distinctions between feelings are easily lost especially at low intensities [21]. Furthermore, physiological responses vary between individuals or even for the same individual depending on the context so it is hard to generalize information from these proxies [19], [25].

III. Biomusic Design

A. Auditory Design

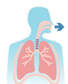
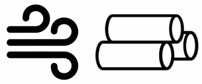
We composed our biomusic using heart rate (HR), breathing rate (BR), and electrodermal activity (EDA) data. Based on Bernardi et al.'s finding that physiological effects of music correspond to changes in tempo [26], we designed our biomusic so that alterations in body physiology affected its *tempo* rather than other musical qualities such as pitch, key, or volume. Our biomusic consisted of the following sounds (See Table I):

1) Kick-Drum: The HR tempo controlled a kick drum that played twice in quick succession, simulating the lub-dub (systole-diastole contractions) of a heart. The tempo occurred at three-quarters of the user's HR and did not precisely align with its real-time beating. We scaled the tempo down based on pilot participants' feedback that an unscaled tempo was too anxiety-provoking and they mainly listened to tempo *changes* which would be unaffected by scaling.

2) Tube Air: The BR controlled a "breathing" sound effect that resembled air passing through a tube or the sound of someone exhaling. This sound's tempo was the same as the participants' BR but was not synchronized to exact real-time inhalation-exhalation patterns.

3) Saturated Bass: Detected peaks in the user's phasic EDA (short-term changes in sweat level known to be related to affective states [27]) triggered a saturated bass sound in a binary manner. This sound was designed to be slightly ominous as pilot participants reported that most EDA spikes were during discussions of negative events like breakups or job losses and an ominous sound would be congruous.

TABLE I
THE BIOMUSIC'S COMPONENTS AND AUDITORY SOUND MAPPINGS. SEE SUPPLEMENTAL MATERIALS FOR AUDIO RECORDINGS

Physiological Input Data	Biomusic Sound Mapping
 Heart rate (HR)*	 Kick Drum at 75% of HR tempo
 Breathing rate (BR)*	 Tube Air at 100% of BR tempo
 Electrodermal activity (EDA)*	 Saturated Bass for phasic EDA peaks

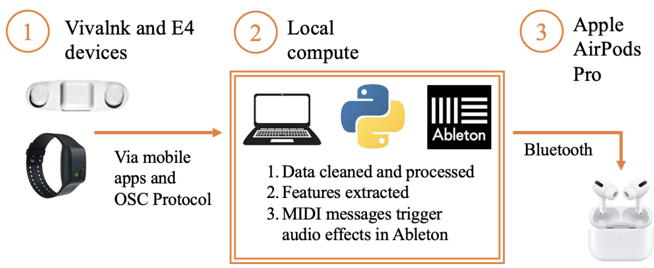


Fig. 1. Biomusic System Architecture. Data flows from wearable devices to local compute that triggers the biomusic sonification. These are played through AirPods Pro worn by a user's conversational partner. Participants were instructed to keep the E4-wearing wrist as still as naturally possible.

B. System Architecture

Physiological data streams were acquired in real-time from an Empatica E4 wristband (4 Hz) and a Vivalnk ECG monitor (2 Hz). These transmitted data to a local laptop (2016 15" Macbook Pro) using the Open Sound Control protocol [28]. The local laptop extracted relevant features and triggered sound effects in Ableton with MIDI messages. This sonification played back to the other user through AirPods Pro which were set to transparency mode to allow the user to participate in a conversation at the same time. See Fig. 1.

1) Data Processing: Data was validated in real time, dropping data outside these ranges and using last value imputation: HR: [45, 120] beats per minute, BR: [8, 30] breaths per minute, EDA: [0, 12] micro-siemens (μS). No further processing was done for HR or BR. EDA movement artifacts were removed if the magnitude of acceleration was greater than 1 in a 2-second rolling window. EDA data was then imputed with a spline, upsampled to 20 Hz, smoothed to a 1-second moving average, run through a Butterworth filter with highcut=3 and order=4, and decomposed into phasic and tonic components using neurokit2's `eda_phasic` function [29]. Finally, neurokit2's `eda_peaks` function detected phasic EDA peaks [29]. Total processing time was tenths of a second.

TABLE II
VALIDATION THAT SYNTHETIC DATA FROM TIMEGAN IS SIMILAR TO REAL DATA VIA A TIME SERIES PREDICTION TASK ON UNSEEN REAL DATA

Prediction Model	R^2	Mean absolute error	Mean root log error
Trained With Real Data	0.901	0.030	0.001
Trained With Synthetic Data	0.890	0.032	0.001

C. Synthetic Biomusic

To test whether the connective qualities of real biomusic are due to the information it encodes or other aspects of its music, we sought to create "sensor-less" biomusic that encoded no real physiological information but was statistically and perceptually indistinguishable from real biomusic.

To achieve this goal, we trained a time-series generative adversarial network (TimeGAN) [8]. TimeGAN is a generative model that captures stepwise conditional distributions from the training data. TimeGAN training involves four networks: encoding and decoding networks map time series features to an embedding space, an adversarial network learns the data's temporal dynamics, and the last network produces next time step outputs. These outputs are trained on supervised loss from the original data, enabling TimeGAN to capture time conditional distribution data. TimeGAN can thus capture changes in distributions over time (e.g., if HR values rise over time because participants are getting tired).

We collected 10 hours of data as training data from 8 pilot participants as they participated in the same exercises as the planned experiment. Our best-performing model was based on the original paper implementation [30], with hidden dims = 24, $\gamma = 1$, noise dims = 32, batch size = 128, learning rate = $5e-4$, and epochs = 10,000.

To validate that our TimeGAN produced synthetic data with the same patterns as real data, we trained one time series prediction model on synthetic TimeGAN physiological data and another identical model on real physiological data. We compared these models on how well they predicted an unseen, *real* physiological time series (Table II) and found they performed equally well, indicating our synthetically generated data was statistically similar to the real data. Further validation was done pre and post-study using dimensionality reduction and visual inspection.

Ultimately, the quality of our synthetic biomusic was validated by our study participants. None of the participants answered that they could discern a difference between the two biomusic conditions.

IV. STUDY METHODS

A. Study Design

For RQ1, we hypothesized that our empathy and closeness metrics would be higher for the real biomusic condition than the synthetic one. We believed the listener would not be able to draw meaningful information from the synthetic biomusic

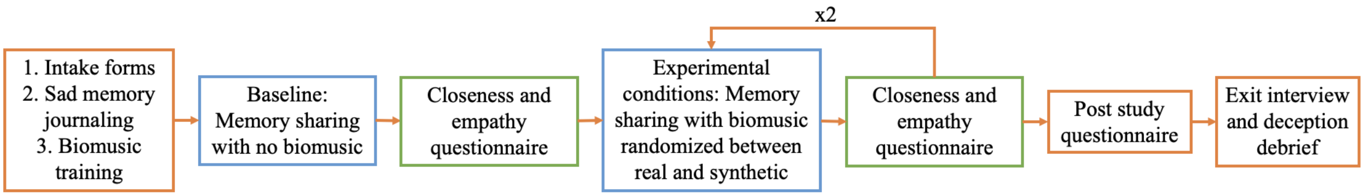


Fig. 2. Overview of the study protocol which took approximately 105 minutes, with memory sharing lasting about 35 minutes.

because it is unrelated to the situation and that its out of sync nature would create an incongruous and distancing effect.

To answer RQ1, we tested the effects of hearing real vs synthetical biomusic on the empathy and closeness between pairs of participants who shared vivid sad memories with each other. Vivid sad memories were chosen to try to provoke active affective physiological responses. We constrained memory sharing to sad memories only to limit the number of variables differing per dyad.

This was a within-subjects experiment, so all participants received all three conditions. In all cases, we measured a no music baseline first and then randomized whether the real biomusic condition came second or third. The no music baseline came first because during pilot studies we noticed if it came after a biomusic condition, participants would guess it to be a "control" and not treat it seriously. Putting it first allowed us to best measure closeness and empathy without biomusic. We did not notice ordering effects because of this.

The study was approved by an internal Institutional Review Board (IRB) review, as well as security and privacy reviews, to ensure participants' data was kept anonymous and secure. Participants received a $100 gift card for their participation.

B. Participants

There were 24 total dyads for a total of 48 participants. 24 identified as female, 22 identified as male, and 2 identified as non-binary / gender diverse. Participants in each dyad did *not* know each other before the study. Binned age data ranged from 22 and 65 with the mean lying in the 36-45 bin. Participants were members of a technology company screened for communication, cognition, or emotional impairments and who had not used a biosignal sharing system before. Our final analysis included 22 dyads; data from two dyads was excluded because of an audio malfunction and non-fluency in English.

C. Procedure

All participants provided informed consent prior to arrival. Upon arrival, they were equipped with an Empatica E4 wristband on their non-dominant wrist and a Vivalnk chest monitor. Participants completed an intake form containing questions on demographics, the Toronto Empathy Questionnaire (TEQ) [31], the Multidimensional Assessment of Interoceptive Awareness (MAIA) [32], the Emotion section of the Goldsmiths Musical Sophistication Index (GMS) [33], and the Emotional Contagion Scale (EC) [34]. These were used to screen for potential issues affecting participants' ability to empathize, emotionally communicate, or interact with biomusic. No issues were found.

Next, participants journaled about three vivid memories involving strong sadness, knowing they would share these memories later. Then they listened to a sample of biomusic and received training on how changes in body physiology affected each biomusic component. The volume was adjusted per participant so they could engage in conversation and listen to biomusic simultaneously with low cognitive load.

The investigator then left the room to observe the participants through a one-way mirror. The study protocol involved three rounds of memory sharing, with randomized memory order and participant sharing order. Each participant had four minutes to share during each round. The listening participant was instructed to respond naturally but to allowing the speaker to share their memory without being too interrupted. The first round served as a baseline measurement without biomusic. The second and third rounds involved both participants randomly receiving either real or synthetic biomusic, with no indication that one of the conditions was synthetic.

After each round, participants completed a short survey assessing empathy, closeness, and information gained from the biomusic. A post-study questionnaire asked about their overall experience using biomusic. This was collected at the end to prevent excessive reflection on the biomusic during the study. Finally, the investigator returned for an exit interview, debriefed the participants, revealed the synthetic biomusic, and distributed their payment. See Fig. 2.

D. Measures

To measure empathy, we utilized seven empathy-related questions developed by Haegerich et al. which assess an individual's ability to comprehend and resonate with another's thoughts and feelings, such as "I can really feel what my partner must have been feeling during this event" [35]. Participants responded using a 7-point scale, and the scores were averaged into an Empathy Composite score.

To evaluate information gained, participants rated the amount of information they obtained from the biomusic on a 7-point scale, comparing it to three channels: body language, facial expressions, and tone of voice.

The post-study questionnaire contained questions assessing participants' interest in listening to others' biomusic, their comfort in sharing their own biomusic, and general qualitative feedback regarding the biomusic and its utility.

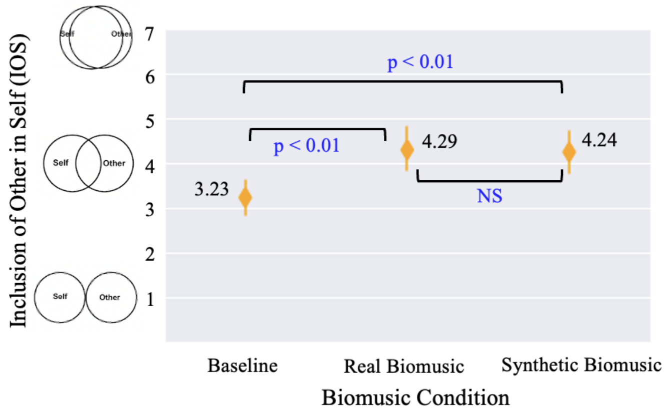


Fig. 3. Effect of biomusic on closeness. There was a significant difference in rated closeness between both biomusic conditions and the baseline condition but not between the synthetic and real biomusic conditions.

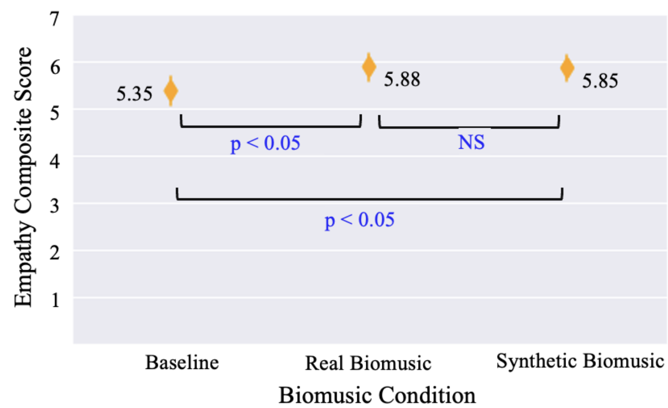


Fig. 5. Effect of biomusic on empathy. There was a significant difference in rated empathy between both biomusic conditions and the baseline condition but not between the synthetic and real biomusic conditions.

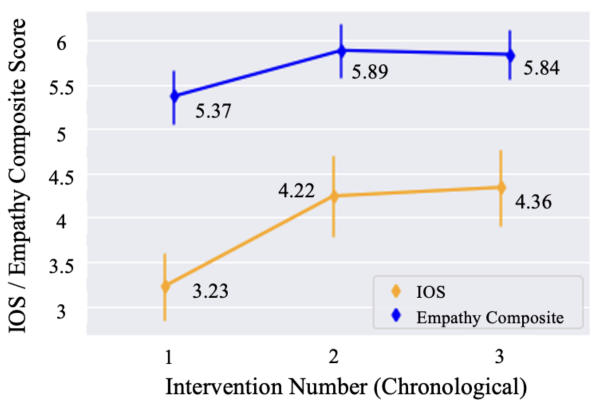


Fig. 4. Effect of intervention ordering. We did not observe an ordering effect where empathy or closeness increased with each intervention.

V. Results

We answer RQ1 by comparing the impacts that synthetic and real biomusic have on closeness (Section V-A) and empathy (Section V-B). We answer RQ2 by showing how informative users found the biomusic (Section V-C). We answer RQ3 by extracting themes from participants' responses on how they related to both synthetic and real biomusic (Section V-D).

A. What is biomusic's impact on closeness?

Fig. 3 shows the average and 95% confidence interval for responses in terms of closeness. Both synthetic and real biomusic were found to increase closeness compared to baseline. Our IOS scores were as follows: Baseline ($M = 3.27, SD = 1.41$), Real Biomusic ($M = 4.29, SD = 1.63$), Synthetic Biomusic ($M = 4.24, SD = 1.57$). Before running a repeated measures ANOVA, we used Levene's test to check for equality of variance, and the result was not significant: $L(2, 120) = 0.35, N.S.$. Our repeated measures ANOVA on these conditions found a significant difference: $F(2, 120) = 5.62, p = 0.004$, Cohen's $d = 0.66$. A post-hoc Tukey test showed the differences between real and baseline and between synthetic and baseline were both significant at $p < 0.006$ and $p = 0.009$. There was no significant difference between real and synthetic biomusic ($p = 0.892$).

Fig. 4 shows the average and 95% confidence interval for closeness scores for each chronological intervention. Scores were lower for the first intervention, which was always the Baseline, but we did not see increases in closeness over all three interventions which would indicate ordering effects. Closeness scores plateaued during interventions 2 and 3, which were randomized between real and synthetic biomusic.

B. What is biomusic's impact on empathy?

Fig. 5 shows the average and standard error responses in terms of empathy. Both real and synthetic biomusic were found to increase empathy compared to the baseline condition. Our Empathy Composite scores were as follows: Baseline ($M = 5.35, SD = 1.10$), Real Biomusic ($M = 5.88, SD = 0.98$), Synthetic Biomusic ($M = 5.85, SD = 0.91$). Before running a repeated measures ANOVA, we used Levene's test to check for equality of variance, and the result was not significant: $L(2, 120) = 1.12, N.S.$ Our repeated measures ANOVA found a significant difference between these conditions: $F(2, 120) = 3.54, p = 0.032$, Cohen's $d = 0.50$. A post-hoc Tukey test showed the differences between real and baseline and between synthetic and baseline were both significant at $p = 0.030$ and $p = 0.040$. There was no significant difference between real and synthetic biomusic ($p = 0.909$).

Fig. 4 shows the average and 95% confidence interval for Empathy Composite scores according to the chronological intervention number. Scores were lower for the first intervention, which was always the Baseline, but we did not see increases in empathy over all three interventions which would indicate ordering effects. Empathy scores plateaued during interventions 2 and 3, which were randomized between real and synthetic biomusic.

C. How much information does biomusic provide?

Fig. 6 shows the average and standard error responses in terms of information gained. Both the real and synthetic

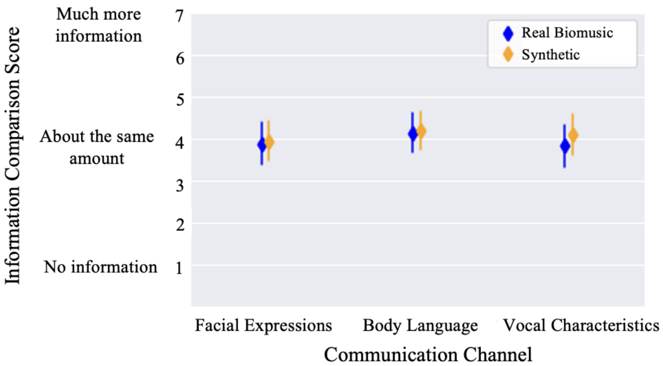


Fig. 6. Information gained from biomusic. Users reported gaining the same amount of information from both biomusics as they did from conventional methods such as observing body language, facial expressions, or voice.

biomusic are rated as providing the same amount of information about a participant's response as traditional social information channels, such as observing their facial expressions or body language or paying attention to their tone of voice. Welch's t-test returned a non-significant result for differences in information gained between synthetic and real biomusic.

D. How do users relate to biomusic?

We present several salient themes for how users related to the biomusic. No users perceived any differences between the synthetic and real biomusic, so their responses applied to their general experience of both biomusics.

1) Biomusic as "The Truth": Without being prompted, 11 of our participants indicated they felt biomusic revealed the *truth* of what they or their partner were feeling. Many participants commented that this made the interaction automatically vulnerable as the biomusic made it so they could not hide what they were feeling.

> *"**Physiological signs show the true emotions**... It was a little uncomfortable, letting someone else understand **the full extent of your emotions**, made me feel vulnerable."* - P43

> *"When I said that my dad and uncle were close, I realized my heart may have sped up because they actually had a fight before he died. I came back and corrected that because **I thought my heartrate may have reflected that wasn't accurate. It was like a lie detector test.**"* - P21

> *"I felt a bit more nervous because I knew there was **no way to hide** or gloss over the **real** emotional impact [of] the events I was describing."* - P11

2) Biomusic provides information they could not have gotten otherwise: Without being prompted, 25 of 44 participants reported biomusic added information, 18 of these 25 said **biomusic provided information they could not have received from traditional sources such as body language or facial expressions**, and only 5 of 44 indicated they did not receive new information. This provided connective benefit especially for those who thought they were paired with less expressive partners.

> *"I got more insight into the emotions behind the story as **her words didn't really tell me what she was feeling**."* - P33

> *"Sometimes facial expressions and body language can be misleading and people have learn[ed] to control those but **[the biomusic] shows the depth of emotions the other person is feeling**."* - P43

> *"It was an interesting way to communicate with them. It felt like **an added layer of communication that I've never had otherwise.**"* - P40

3) Biomusic provides emotional validation and encourages prosocial behavior: The bidirectional nature of our biomusic allowed some users to draw the validating conclusion that their partners were emotionally affected.

> *"I was willing to share more since I knew they were feeling the same emotions that I was feeling and understood the sadness I was feeling."* - P43

> *"I remember that her heart rate was picking up during the story I was telling so **I knew she was invested**... **I thought it was validating** since it meant she was listening."* - P34

> *"I really enjoyed getting additional cues about her emotions and what she was going through. She also reacted strongly when I had an intense emotional story to tell, which I appreciated because **I knew she was resonating with me**."* - P3

The biomusic also caused participants to be more empathetic (more understanding, and prosocial [9]).

> *"When the heartbeat increased I feel like my active listening kicked in a bit higher as **I knew it was a tough part of the story for her so I wanted to hear it all well** and not miss any part or it nor say/do anything that would make her feel worse."* - P39

4) Heart rate was the most useful component: 39 participants mentioned using the HR sound, 11 mentioned using the EDA peak sound, and 5 mentioned using the BR sound. 24 of the 39 who got value from the HR sound said they listened for when its tempo was increasing.

VI. Discussion

A. Lack of Difference between Real and Synthetic Biomusic

Both real and synthetic biomusic increased our Empathy Composite and IOS metrics compared to baseline, and we found no significant difference between them. This surprised us as the synthetic biomusic was incongruous and unsynced with a user's real physiology. Using Kuan et al.'s measure of positive physiological linkage in emotional settings, we found that the linkage between our synthetic and real biosignals was just 0.544, where 1 is perfect linkage and 0 is no linkage at all, meaning we had about a random amount of linkage [36]. Previous research has demonstrated that the connecting effects of biomusic depend upon its congruence with other affective signals (e.g., facial expressions) [6]. Incongruent biomusic results in lower ratings of emotional congruence and changes emotion perception, so we expected our random linkage would lower our empathy and closeness ratings.

However, most participants were surprised to learn that one of the sharing rounds had synthetic biomusic. They were often sure that their partners had been responding to their story. We believe this is because physiological signals as proxies for affect are highly (mis)interpretable. Just as two different news channels can cover the same event in opposite ways, we observed participants interpret changing physiological signals as evidence for whatever reactions they would normally expect, especially if primed by other information channels like facial expressions. Rising *and* falling heart rate both meant someone was attentive. Users are also largely unfamiliar with physiological signals and perhaps more training is needed to develop instincts for incongruous information channels.

But incongruence may be irrelevant. Laurenceau et al. find that partner emotional disclosure and perceived partner responsiveness are important factors for building intimacy in a relationship [37]. These qualities are provided by both biomusics. Users rated both biomusics as giving them significant amounts of information. Their responses indicate biomusic creates high levels of emotional vulnerability, emotional validation, and prosocial behavior. We conclude that the overall effect is that the vulnerability and validation created and information communicated allow users to perceive more emotional disclosure and partner responsiveness. The increase in perceived disclosure and responsiveness is underscored in responses about non-expressive partners. We hypothesize these factors outweigh the incongruousness of synthetic biomusic, leading to the observed statistically significant increases in the IOS ratings and Empathy Composite scores for both biomusics compared to baseline. It seems users prefer *any* response to none at all. Indeed, Laurenceau et al. find that **emotional** self-disclosure is more important for creating intimacy than **factual** self-disclosures [37].

It seems likely to us that these effects would strengthen if biomusic were used in non-in-person settings, such as in VR or over a voice call where traditional sources of information are limited or even absent.

B. Mitigating Over-Indexing on Biomusic

Participants may have believed biomusic revealed the "truth" about how they or their partner were feeling because our sound design included realistic sounding heartbeats. Although users knew these were sound effects, some participants commented that it felt as though they had a "superpower" and were inside the others' body and were able to listen to their heart. As heartbeats were the most listened to component, we believe using a less-realistic heartbeat such as a cartoon-like one could decrease this effect.

Additionally, a longer training period with biomusic or any biosignal sharing system would help users understand the natural fluctuations in body physiology that occur and read into their changes less. We also believe it is important that users are trained to understand the limitations of physiological signals as proxies for affect as discussed in Section II-D.

VII. Limitations and Future Work

This laboratory study examined biomusic usage in a controlled setting with participants who do not know each other. Our findings' generalizability is limited because our participant pool was employees from a US technology company.

Ordering effects may have impacted comparisons with our baseline which always preceded other conditions. We did not observe strong evidence for this in Fig. 4 as scores were similar between Interventions 2 and 3, suggesting that lower scores for Intervention 1 were due to its Baseline nature. Comparisons between the biomusic conditions, our research's main focus, were counterbalanced and less likely to be affected.

Our study's major limitation is the exclusive focus on sad memories, restricting the range of affective experiences and biomusic encoding. Future research should explore broader affective responses. Participants may also have restricted their sharing around these vulnerable memories. They found biomusic helpful for interacting with less expressive individuals, suggesting potential applications for its use in contexts with limited facial, bodily, or vocal expression, such as virtual reality or the metaverse.

Both synthetic and real biomusic demonstrated similar positive effects on empathy and closeness, but further research is needed to understand these findings. The effect of users' awareness on whether biomusic is synthetic and the empathic effects of biomusic sound design should be studied further.

VIII. Conclusion

As loneliness has become a larger and larger problem in our society, it has become increasingly important to find ways of improving our ability to connect with one another. This study finds evidence that biomusic can be an effective method for doing so as biomusic is shown to increase empathy and closeness in in-person dyad interactions. We also find that biomusic provides as much information to users as traditional social information channels such as body language, suggesting that it could be a rich medium for creating connection, especially if users were trained to understand its nuances. Interesting, we fail to find differences in the effects of real vs synthetic biomusic; the empathic benefits of real biomusic are also seen in synthetic biomusic despite synthetic biomusic being unrelated to a user's true physiology. This opens up the possibility of synthetic biomusic as an empathic technology, with exciting applications for using it to augment interactions where real biomusic would not be possible to implement, such as in sensorless environments or virtual interactions.

IX. Ethical Impact Statement

Information on IRB, informed consent, and compensation is in Section IV. Data was anonymized with unlinked identifiers. Participants were given full control over their data retention.

We note our study may not generalize in Section VII. Similarly, our TimeGAN model training was ungeneralizable, and will not be used further. Future work should be aware of cultural connotations in biomusic design. Our sonification was realistic sounding, so we think connotations were limited.

Societal concerns about include contributing to the ability to fake realistic physiological responses. Unintended effects could be erosion of trust in biosignal sharing or decreasing human-to-human interactions by making virtual avatars with biomusic more acceptable to people. Research to prevent this should include "fingerprinting" synthetic biomusic. Societal benefits include potentially increasing the ability of people to connect, especially those who are unable to use traditional social information channels. To mitigate risks, further research should consult these groups on biomusic design.

Acknowledgment

Thank you to Hernan Asorey for his willingness to create opportunities for those around him.

References

[1] D. A. Cox, "The state of american friendship: Change, challenges, and loss," Apr 2022. [Online]. Available: https://www.americansurveycenter.org/research/the-state-of-american-friendship-change-challenges-and-loss/

[2] J. T. Cacioppo and S. Cacioppo, "The growing problem of loneliness," *The Lancet*, vol. 391, no. 10119, p. 426, 2018.

[3] J. H. Janssen, J. N. Bailenson, W. A. IJsselsteijn, and J. H. Westerink, "Intimate heartbeats: Opportunities for affective communication technology," *IEEE Transactions on Affective Computing*, vol. 1, no. 2, pp. 72–80, 2010.

[4] M. Hassib, D. Buschek, P. W. Wozniak, and F. Alt, "Heartchat: Heart rate augmented mobile chat to support empathy and awareness," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 2239–2251.

[5] P. Slovák, J. Janssen, and G. Fitzpatrick, "Understanding heart rate sharing: towards unpacking physiosocial space," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2012, pp. 859–868.

[6] R. M. Winters, B. N. Walker, and G. Leslie, "Can you hear my heartbeat?: Hearing an expressive biosignal elicits empathy," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–11.

[7] N. Howell, L. Devendorf, R. Tian, T. Vega Galvez, N.-W. Gong, I. Poupyrev, E. Paulos, and K. Ryokai, "Biosignals as social cues: Ambiguity and emotional interpretation in social displays of skin conductance," in *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, 2016, pp. 865–870.

[8] J. Yoon, D. Jarrett, and M. Van der Schaar, "Time-series generative adversarial networks," *Advances in neural information processing systems*, vol. 32, 2019.

[9] P. Oliveira-Silva and Ó. F. Gonçalves, "Responding empathically: A question of heart, not a question of skin," *Applied psychophysiology and biofeedback*, vol. 36, no. 3, pp. 201–207, 2011.

[10] A. Aron, E. N. Aron, and D. Smollan, "Inclusion of other in the self scale and the structure of interpersonal closeness." *Journal of personality and social psychology*, vol. 63, no. 4, p. 596, 1992.

[11] C. George and M. Hassib, "Towards augmenting ivr communication with physiological sensing data," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–6.

[12] F. Liu, "Expressive biosignals: Authentic social cues for social connection," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–5.

[13] F. Liu, G. Kaufman, and L. Dabbish, "The effect of expressive biosignals on empathy and closeness for a stigmatized group member," *Proceedings of the ACM on Human-Computer Interaction*, vol. 3, no. CSCW, pp. 1–17, 2019.

[14] S. Cheung, E. Han, A. Kushki, E. Anagnostou, and E. Biddiss, "Biomusic: An auditory interface for detecting physiological indicators of anxiety in children," *Frontiers in neuroscience*, vol. 10, p. 401, 2016.

[15] J. Frey, M. Grabli, R. Slyper, and J. R. Cauchard, "Breeze: Sharing biofeedback through wearable technologies," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–12.

[16] F. Grond, M. A. Cascio, R. Motta-Ochoa, T. Tembeck, D. Ten Veen, and S. Blain-Moraes, "Participatory design of biomusic with users on the autism spectrum," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2019, pp. 1–7.

[17] S. Blain-Moraes, S. Chesser, S. Kingsnorth, P. McKeever, and E. Biddiss, "Biomusic: A novel technology for revealing the personhood of people with profound multiple disabilities," *Augmentative and Alternative Communication*, vol. 29, no. 2, pp. 159–173, 2013.

[18] R. W. Picard, *Affective computing*. MIT press, 2000.

[19] M. M. Bradley and P. J. Lang, "Measuring emotion: Behavior, feeling, and physiology." 2000.

[20] P. J. Lang, "The emotion probe: Studies of motivation and attention." *American psychologist*, vol. 50, no. 5, p. 372, 1995.

[21] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *science*, vol. 221, no. 4616, pp. 1208–1210, 1983.

[22] J. Hernandez, I. Riobo, A. Rozga, G. D. Abowd, and R. W. Picard, "Using electrodermal activity to recognize ease of engagement in children during social interactions," in *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing*, 2014, pp. 307–317.

[23] S. Shahrestani, E. M. Stewart, D. S. Quintana, I. B. Hickie, and A. J. Guastella, "Heart rate variability during adolescent and adult social interactions: A meta-analysis," *Biological psychology*, vol. 105, pp. 43–50, 2015.

[24] J. A. Haythornthwaite, D. E. Anderson, and L. H. Moore, "Social and behavioral factors associated with episodes of inhibitory breathing," *Journal of behavioral medicine*, vol. 15, no. 6, pp. 573–588, 1992.

[25] G. G. Berntson, K. S. Quigley, G. J. Norman, and D. L. Lozano, "Cardiovascular psychophysiology." 2017.

[26] L. Bernardi, C. Porta, G. Casucci, R. Balsamo, N. F. Bernardi, R. Fogari, and P. Sleight, "Dynamic interactions between musical, cardiovascular, and cerebral rhythms in humans," *Circulation*, vol. 119, no. 25, pp. 3171–3180, 2009.

[27] G. Geršak, "Electrodermal activity-a beginner's guide," *Elektrotehniski vestnik*, vol. 87, no. 4, pp. 175–182, 2020.

[28] M. Wright, A. Freed, and A. Momeni, "2003: Opensound control: State of the art 2003," *A NIME Reader*, pp. 125–145, 2017.

[29] D. Makowski, T. Pham, Z. J. Lau, J. C. Brammer, F. Lespinasse, H. Pham, C. Schölzel, and S. H. A. Chen, "NeuroKit2: A Python toolbox for neurophysiological signal processing," *Behavior Research Methods*, vol. 53, no. 4, pp. 1689–1696, 2 2021.

[30] J. Yoon, "Timegan," https://github.com/jsyoon0823/TimeGAN, 2021.

[31] R. N. Spreng*, M. C. McKinnon*, R. A. Mar, and B. Levine, "The toronto empathy questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures," *Journal of personality assessment*, vol. 91, no. 1, pp. 62–71, 2009.

[32] W. E. Mehling, M. Acree, A. Stewart, J. Silas, and A. Jones, "The multidimensional assessment of interoceptive awareness, version 2 (maia-2)," *PloS one*, vol. 13, no. 12, p. e0208034, 2018.

[33] D. Müllensiefen, B. Gingras, J. Musil, and L. Stewart, "The musicality of non-musicians: An index for assessing musical sophistication in the general population," *PloS one*, vol. 9, no. 2, p. e89642, 2014.

[34] R. W. Doherty, "The emotional contagion scale: A measure of individual differences," *Journal of nonverbal Behavior*, vol. 21, no. 2, pp. 131–154, 1997.

[35] T. M. Haegerich and B. L. Bottoms, "Empathy and jurors' decisions in patricide trials involving child sexual assault allegations," *Law and human behavior*, vol. 24, no. 4, pp. 421–448, 2000.

[36] K.-H. Chen, C. L. Brown, J. L. Wells, E. S. Rothwell, M. C. Otero, R. W. Levenson, and B. L. Fredrickson, "Physiological linkage during shared positive and shared negative emotion." *Journal of Personality and Social Psychology*, vol. 121, no. 5, p. 1029, 2021.

[37] J.-P. Laurenceau, L. F. Barrett, and P. R. Pietromonaco, "Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges." *Journal of personality and social psychology*, vol. 74, no. 5, p. 1238, 1998.