Physiological synchronization: does it influence the empathy for one another?

Derk H. J. Krukkert - 0957537

Wei-Hsi Lin - 1300881

Pranav Muppirishetty - 1331108

Dominique M. E. Peeters - 1394762

Supervisors:

prof.dr. J.H.D.M. Westerink

MSc M. A. Feijt

Department of Industrial Engineering & Innovation Sciences, Eindhoven University of Technology

0HM170: HTI Research Project

Group 5 - Are we ‘in sync’?

February 1, 2020

Abstract

This research provides an insight into the relationship between physiological synchrony (PS) and empathy, as well as the influence of presenting PS as feedback on empathy. Using a combination of experimental procedures used in previous studies, this study attempts to resolve inconsistent findings regarding PS-empathy relationship. Synchronization of Heart Rate (HR) were measured in a between-dyad design. The PS was measured while the participants were subjected to an emotion-evoking video either with or without real-time feedback in the form of a colored border around the video. Empathy and social connectedness were measured by a self-reported questionnaire. Regression on data from 105 participants show a significant positive correlation between PS and empathy for participant dyads that were in the feedback condition and that did not know each other ($r(47) = .41$, $p < .01$), whereas the correlation that includes participants in both feedback and no-feedback condition is not statistically significant. Another moderated regression of empathy on PS with PS feedback as the moderator indicates that feedback has a negative correlation with empathy ($B_2 = -0.20$, $t(104) = -2.44$, $p = .02$), and it has a significant moderating effect ($B_3 = 0.86$, $t(104) = 2.77$, $p = .01$). Additionally, PS score significantly predicted empathy in the feedback condition ($B_1 = 0.70$, $t(48) = 3.04$, $p < .01$). Results from this study can be implicated in computer-mediated communication, where providing PS feedback can lead to an increased empathy and possibly resolve a part of the lack of non-verbal cues.

Content

1. Introduction

Face to face communication normally entails the understanding and use of non-verbal cues like eye contact, facial expressions, gestures and interpersonal distance. Online communication lacks these elements, making the conversation devoid of completeness according to Carter (2003). Carter (2003) focused on the fact that people use Computer-Mediated Communication more and more for information exchange, to meet new people and to bond with others. Since online communication lacks non-verbal cues, individual differences in interpretation can often lead to miscommunication. For instance, people can perceive someone as rather impolite, even when that was not the intention. Since society at large increasingly depends on online communication, new ways in which this interaction can be improved are being explored. One such way entails the use of physiological information of the conversation partners to improve the quality of online communication, specifically with regards to perceived empathy. The current study involves using information on the synchrony of physiological signals (e.g., Heart Rate (HR)) of two persons to improve perceived empathy and social connectedness. In the next paragraphs, first, the concept of empathy and its importance in online communication will be discussed. Then, the potential benefits of physiological signals is presented and the concept of physiological synchronization is introduced. In the remainder of the introduction, two previous studies on this topic with conflicting findings will be discussed. The focal point in this section will be the key issues and questions that arise from these studies. This also introduces the objectives of the current study. Finally, the specific research question and hypotheses of the current study are explained in detail.

The concept of empathy is very important in social interactions. From an evolutionary perspective, empathy increases the chances of survival by aiding social support (Preston & de Waal, 2002). It is a major determinant for prosocial and altruistic behavior (Batson, Duncan, Ackerman, Buckley, & Birch, 1981; Hoffman, 1984). Furthermore, empathy is positively related to relationship satisfaction (Cramer & Jowett, 2010). Moreover, empathy is negatively related to aggressive and antisocial behavior (Miller & Eisenberg, 1988). Existing literature consists of a large number of different definitions, making it a difficult concept to grasp. Cuff, Brown, Taylor, and Howat (2016) performed a critical review of a range of definitions of empathy to develop a new conceptualisation that is more consistent. Their extensive research is summarized in the following definition:

> "*Empathy is an emotional response (affective), dependent upon the interaction between trait capacities and state influences. Empathic processes are automatically elicited but are also shaped by top-down control processes. The resulting emotion is similar to one's perception (directly experienced or imagined) and understanding (cognitive empathy) of the stimulus emotion, with recognition that the source of the emotion is not one's own.*" (p. 150).

Another concept that is related to empathy is social connectedness. Van Bel, IJsselsteijn, and De Kort (2008) explain this term as having a sense of belonging due to a solid amount of social contacts. It is related to empathy in the aspect that it is also an important factor in social interaction.

According to Preece (1999), empathy is to be perceived less in Computer-Mediated Communication, because there are less non-verbal cues available. Physiological information can be a potential solution to resolve this deficit in perception of empathy. First, it is important to know how physiological information is related to interpersonal communication. Janssen, Bailenson, Ijsselsteijn, and Westerink (2010) found that receiving a conversational partner's physiological information affects the psychological state of the participant. Through a virtual reality based experiment, the study found that the participants tend to increase the physical distance between themselves and the confederate when presented with the confederate's fabricated heartbeat. It was inferred that a compensation mechanism came into effect because the physiological information is interpreted as an intimate non-verbal cue. The advantage of using physiological signals for communication is that they are implicitly available and do not need the user's conscious effort into formulating the content of expression.

Synchronization of actions, gestures or behaviors among a group of people is a common phenomenon, as can be derived from various cultures and groups of people (Wiltermuth & Heath, 2009). It is often bodily, vocal and postural mimicry of nature (Hatfield, Rapson, Ye, Decety, & Ickes, 2009): This can be mimicking facial expressions or posture by taking a seat and crossing the same leg. While the external physical manifestation of synchrony is intriguing in itself, it also manifests at a much deeper level; Human physiology has been shown to exhibit substantial synchrony between persons (DiMascio, Boyd, & Greenblatt, 1957; Levenson & Gottman, 1983). Palumbo et al. (2016, p. 100) define Physiological Synchrony (PS) as "any interdependent or associated activity identified in the physiological processes of two or more persons." In a study conducted by Marci, Ham, Moran, and Orr (2007), a significant positive correlation was found between skin conductance concordance and patient ratings of perceived therapist empathy. This relation was found again in a study by Messina et al. (2013). Their experiment involved a fabricated therapist-patient role-playing setting where volunteer students pretended to be a patient and had to discuss a personal problem with a "listener" that was either a real therapist, a psychologist or a non-therapist. These studies were an investigation of the notion that the quality and nature of interaction shared by the participants has an underlying shared physiological change. Even though extensive studies have been carried out on PS, not much research has been done that focuses on the role of communicating the measured PS to the two involved persons, consequently using PS in a feedback-role. This PS feedback could possibly be used to fill in the gap between two communicating people, when no other verbal cues are available.

Building on the research findings of Marci et al. (2007), Okel (2018) investigated the effect of explicitly displaying the level of PS between two persons in a lab setting on perceived empathy and social connectedness. The researcher found a positive relation between the level of PS feedback and perceived empathy, whereas only partial support was found that indicates a positive relation between the level of PS and social connectedness. This indicates that showing a higher level of PS results in a higher perceived empathy between the two involved persons. However, this study used simulated PS feedback and the participants were randomly assigned to two conditions: pre-set feedback of high or low PS. Additionally, this study used a confederate that played the role of a participant in each session. Put together, these PS feedback were simulated and a confederate was used, hence these results are not generalizable. This raises the question whether high feedback will still result in higher empathy, or if there is a relation at all between showing this PS feedback and the perceived empathy, when measuring both participants in real-time and communicating this measurement to them.

In response to Okel (2018), Van Laar (2019) investigated if these effects would indeed be present when the level of PS was actually measured in real-time instead of using simulated PS levels. Besides the effect of PS on perceived empathy and social connectedness, this researcher was also interested in the effect of PS on team performance and social flow. Moreover, he also examined whether the correlations between PS and each of the other four variables mentioned above differed between the three conditions: showing a high PS, showing a low PS, and not showing the PS at all (i.e., by comparing the correlation slopes of all conditions). In contrast to Marci et al. (2007) and Messina et al. (2013), Van Laar (2019) used HR as a measure for PS. Interestingly, this study found no significant correlations between PS and each of the following variables: perceived empathy, social connectedness, and team performance. Likewise, the correlations between PS feedback and the aforementioned variables were not statistically significant. This latter finding by Van Laar is inconsistent with earlier findings by Okel (2018). Furthermore, Van Laar found no evidence that suggests that PS feedback has significant moderating effects on the three variables mentioned above.

To explain this surprising finding, Van Laar (2019) discusses several possible explanations, mainly regarding the difference in experimental setup between his study and the one from Okel (2018) (e.g., real PS feedback vs simulated PS feedback). Of the three possible explanations mentioned by Van Laar (2019), one possible explanation in particular is relevant to the current study: Van Laar’s experiment included a different experimental task compared to Okel’s (2018) study, where dyads of participants watched an emotionally salient video. The experimental task by Van Laar, on the other hand, involved a task aimed at team performance, namely the Desert Survival Task and the Lost at Sea Task, in which the two participants in a dyad had to discuss certain matters with each other. Since this researcher was also interested in the effect of PS on team performance, this task arguably made sense for his particular study. But these discussion tasks were lacking emotional variation and could therefore potentially explain the conflicting findings regarding the relation of PS with perceived empathy and social connectedness. Furthermore, Okel (2018) used skin conductance as a measure for PS and Van Laar (2019) used HR. Since these are two different measures, it should be taken into account that this can influence the way of how PS is interpreted and how PS feedback is related to perceived empathy.

Consequently, the current study is a combination of elements of the experimental setup of Okel (2018) with that of Van Laar (2019), which will be explained in more detail in the subsequent sections. Therefore, the research question of this study is:

What is the relation of PS with perceived empathy, and how is this effect moderated by showing vs. not showing PS feedback?

The answer to this may shed light on the question why Van Laar's study failed to corroborate initial findings by Okel. Moreover, it will facilitate a deeper understanding of PS in relation with these elements.

As discussed above, previous research by Marci et al. (2007) found a significant positive correlation between skin conductance concordance and patient ratings of perceived therapist empathy. Similarly, Okel (2018) found that feedback of PS is positively related to perceived empathy. In accordance with these previous findings, the expectation is that PS is positively related to perceived empathy. Hence, the first hypothesis is defined as follows:

H1: PS is positively related to perceived empathy.

According to Janssen et al. (2010), the perception of heartbeat attributed to a conversational partner influences social behavior in a similar manner as traditional non-verbal intimate signals such as gaze and interpersonal distance. This suggests that perceiving information about another person's physiology influences one's judgment and behavior. Accordingly, as discussed above, the study by Okel (2018) demonstrated that dyads with high and low PS feedback scored high and low, respectively, on perceived empathy. The findings of both studies mentioned above suggest that showing a participant dyad, their level of PS, will influence their perceived empathy towards one another. Then, a dyad with a high level of PS, and thus high perceived empathy, will perceive the knowledge about their PS being high, which results in an even higher perceived empathy. In contrast, when a dyad is informed about their PS being low, the perceived empathy, which is already low as a result of low PS, is likely to decrease even further. In terms of a visual representation, this effect could be interpreted as multiple correlation graphs with varying slopes: High PS feedback would yield a steeper slope (i.e., compared to no PS feedback), whereas low PS feedback would yield in a flatter slope. This train of thought can be compressed into the second hypothesis:

H2: The slope of the regression line for feedback condition is positive, and steeper than the slope of the regression line for no-feedback condition .

The next section describes the method that is used to answer the research question. After the method section the results are presented, followed by the discussion and finally the conclusion.

2. Method

2.1 Design

The design of this study involves a lab experiment with a between-dyads design (i.e., one dyad consists of a pair of 2 participants). Participants are randomly assigned to one of two conditions: In the feedback condition, the level of PS between the participants is presented to the participants while watching an emotion-evoking video. In the no-feedback condition, the level of PS is not presented. In both conditions, participants watch the same emotion-evoking video. Afterwards, they fill in two questionnaires on feelings of empathy towards the other participant (Okel, 2018; Van Laar, 2019), and social connectedness (Okel, 2018; Van Laar, 2019). The analysis of the social connectedness variable can be found in the Appendix A, as the primary focus of this study is on perceived empathy. The dependent variables in this study are PS and perceived empathy. The independent variable is the presence or absence of feedback of the real-time PS to the dyad (i.e., feedback or no feedback).

2.2 Participants

For the calculation of the sample size, the focus was put on the determination of the relationship between PS and perceived empathy, since no literature was found on the differences between regression slopes for showing and not showing the PS feedback. Similar to Van Laar (2019), the power analysis for this study was based on previous studies of Marci et al. (2007) and Messina et al. (2013). Both studies focused on the correlation between PS and empathy. In both cases, PS was calculated by participants' Galvanic Skin Response (GSR). It is important to note that there are no studies found linking actual PS in HR to empathy.

A significant correlation of r =.47 (p = .03) was found in the study by Marci et al. (2007). Likewise, Messina et al. (2013) found significant correlations of r = .32 (p = .048) and r = .37 (p = .022). Because the nature of the experimental tasks of the study of Van Laar (2019) is very different from this study, the results of that study are not used for this power calculation. Using these results as a guideline, the correlation used for the power calculation is r = .35. G*Power (http://www.gpower.hhu.de/) was used to calculate for the power analysis. Assuming alpha = .05 and power = .90, this calculation resulted in a sample size of 68 dyads, which means 136 participants were required for this study when performing a correlation test. The output of G*Power can be found in Appendix B. Participants with a severe physical disability, cardiovascular illnesses or clinically impaired emotional perception were excluded from participation to avoid extraordinary data, which could lead to drawing unreliable conclusions.

2.3 Setting and stimulus materials

The setup of the experiment required for the participants to be seated in two separate rooms, room A and room B. A detailed schematic showing the set-up can be seen in Figure 1. The two rooms were identical and each of the rooms contained a desk with a screen, a Mobi device, and a pair of headphones. The Mobi devices were connected to the Mobi laptop in the control room via bluetooth. The Phylo (software to acquire physiological data from Mobi devices; Boschman, 2017) and HRSync (software to calculate PS; Boschman, 2018) applications were run on the Mobi laptop. The realtime PS score calculated by the HRSync application was sent to the processor in participant room A through LAN. The HRSyncColor application (used in the Van Laar study(2019)) was run on the processor in room A to display the colored border in case of the feedback condition, which will be explained shortly. The audio-visual stimulus (along with the colored border in feedback condition) was streamed simultaneously on both the screens in participant rooms A and B through a local VGA connection. The video and the HRSyncColor application were run on the processor in room A through team-viewer application being controlled by the control PC in the control room.

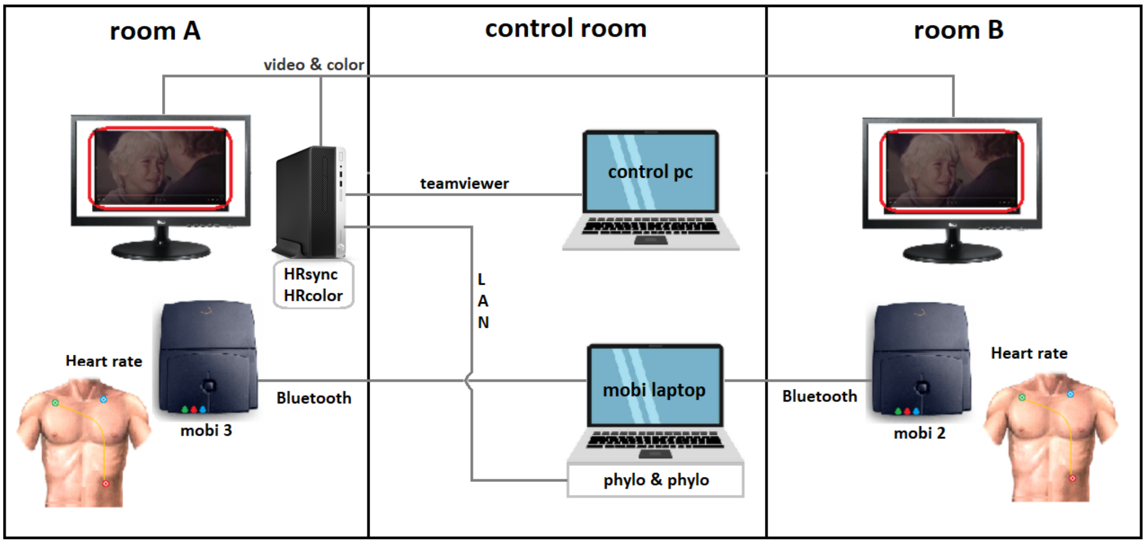

Figure 1. Set-up of the experiment.

For this experiment, an emotion-evoking video was used. This video consists of a small sample from the movie ‘The Champ’ (1979), in which a kid can be seen crying over the death of his father. The same fragment was used in the study by Okel (2018). The content of the video was suitable for this research because a previous study by Gross and Levenson

(1995) showed it was evoking emotions and Van den Broek and Westerink (2009) have found that it influences HR as well.

The feedback condition in this experiment involved showing a colored border around the emotion-evoking video by using the HRSyncColor application. The HRSyncColor application was a simple program that converted the PS values received from HRsync to the colored border. When a high PS was measured between the two participants (e.g., a correlation larger than .2), a green border appeared around the video. The same was done for a low PS (e.g., a correlation below .1), but then a red border was shown. The yellow border appeared when the PS was not high nor low (e.g., correlations between .1 and .2). The color displayed was based on the real-time PS value calculated on a 10 second window. The thresholds that determine when the color should change were determined based on the PS values of 2 dyads of the pilot study and the first 14 dyads in the study. Showing the PS by this colored border was originally implemented in Okel's (2018) experiment. The latter study also included a pilot study to determine the best way of showing the PS and concluded that showing a graph, a number, or a percentage on a bar cost too much cognitive effort. Colors are more efficient and salient. Hence, this study also used a colored border to give PS feedback.

It was also important to know when HR is responding due to the emotion-evoking video. Literature shows that after approximately 4 - 5 seconds, the sympathetic activation is visible in HR signal (i.e., HR is increasing; Appelhans & Luecken, 2006; Ménard, Richard, Hamdi, Daucé, & Yamaguchi, 2015; Smith & Strawbridge, 1968). Therefore the feedback of PS was shown after 5 seconds from the beginning of the emotion-evoking video.

2.4 Measurements

2.4.1 Physiological Synchrony (PS)

The physiological measurement that was used in this study is HR. The HR of each participant was measured with three electrodes connected to a Mobi device. The electrodes were placed using the standard II lead positioning: two electrodes for the bipolar input and one electrode was used as ground. Figure 2 gives a view of this ECG measurement.

Figure 2. ECG measurement using three electrodes.

The PS was calculated per second, using the inter-beat-intervals (IBI) within a custom application HRSync, created by the HTI group of Eindhoven University of Technology (Boschman, 2018; Van Laar, 2019). The PS value of the two participants is defined as the correlation of their HR calculated from IBIs in a window size corresponding to 10 seconds (see Formula 1).

$$r_t = \frac{\sum(IBI_{1k} * IBI_{2k}) - \frac{1}{N} * \sum IBI_{1k} * \sum IBI_{2k}}{\sqrt{\left(\sum IBI_{1k}^2 - \frac{1}{N} * \left(\sum IBI_{1k}\right)^2\right) * \left(\sum IBI_{2k}^2 - \frac{1}{N} * \left(\sum IBI_{2k}\right)^2\right)}}, \ldots \quad \sum \equiv \sum_{k=t-N+1}^{t}$$

Formula 1. Formula for calculation of PS.

IBI_{1k} refers to the inter-beat-interval of subject 1 sampled at the moment indicated by the index *k*. N refers to the window size of calculation, i.e., the number of IBI's taken into the calculation for a single correlation measurement. In the current study, 10 values corresponding to a timeframe of 10 seconds were considered for each correlation measurement in this study). Furthermore, outlier detection and correction were handled in real time by the use of an impulse filter and a median operator implemented in HRSync. Only the data during the experimental video was used for analysis.

2.4.2 Baseline

The baseline measurement was done for one minute, in accordance with Okel (2018). This measurement task involved watching a relaxing video of an aquarium, to make sure that the HR of the participants dropped to their resting HR.

2.4.3 Pilot study

A pilot study was conducted to validate that the emotion-evoking video actually influences HR. Two dyads were included in this pilot study. They first looked at a relaxing video (baseline measurement), and afterwards the experimental video started. Results indicated that the HR of the participants changed after the emotion-evoking video started, so this means HR is affected by the emotion-evoking video. This is in line with the previously mentioned studies (Gross & Levenson, 1995; Van den Broek & Westerink, 2009). Because the relaxing video and the experimental video were played right after each other, it was necessary for the participants to press a marker button on the Mobi device to indicate when each video started and ended. Therefore, prompts were given within the video about when to press this button. According to feedback from the pilot study, it was suggested to clarify the instructions and show them for a longer period of time. Also the background color was adjusted to a greyish color instead of black color, in order to make this background more neutral.

The threshold values indicating different levels of PS were determined from the data of the pilot study and, in addition, the data of the first 14 dyads of the study were taken into account as well. To accomplish this, these first 14 dyads were assigned to the no-feedback condition, in order to be able to calculate threshold values without the participants being influenced by the feedback.

For PS feedback, the green color indicates PS is high, red represent a low PS value, and yellow refers to a medium PS value. The threshold values were set to allow for some of the participants see 60% green 30% red, and some of them see 60% red 30% green. To achieve this, the first and third quartile of all the synchrony values collected from the 2 pilot dyads and first 14 dyads from the experiment were calculated, which were -0.22 and 0.42. The percentages of red and green seen by participants in each high synchronized dyad were 14% and 38% in this case, and the number of high synchronized dyads was only 6 among the 16 dyads. The result was far from what was expected, and the negative threshold values were not reasonable. Further, we compared the threshold values derived from the pilot with the threshold values used in Van Laar’s (2019) study, which were 0.1 and 0.2. Using Van Laar’s (2019) threshold values, the result showed that 58% green and 34% red was seen on average by the high synchronized groups, and low synchronized groups saw about 33% green and 57% red. Since Van Laar’s (2019) threshold values meet the 60-30 criteria better, the final threshold was set to 0.1 and 0.2.

2.4.4 Questionnaires

First, some demographic questions were asked regarding age and gender. To measure the perceived empathy and the perceived social connectedness, two different questionnaires were used. Empathy was measured using 7 questions that have already been used in the study of Okel (2018). The questions were based on earlier research by Plank, Minton, and Reid (1996), but they were transformed to meet the context of this experiment. The questionnaire contained statements about, for instance, if the other person understood the other participant’s feelings in the situation, and if they feel that they are on the same wavelength as the other participant. The list of questions has an internal validity of α = .84 (Van Laar, 2019), which is considered sufficient (see Appendix C for the full questionnaire). The statements were answered on a 5-point Likert scale, where 1 means strongly disagree and 5 means strongly agree.

The next self-report measure indicated social connectedness. The Inclusion of Other in Self (IOS; Aron, Aron & Smollan, 1992) is a single pictorial measure and often used to see how much you feel connected to the other person. Figure 3 shows the options to choose from. The more the “self circle” is placed within the “other circle”, the more they felt connected to the other person.

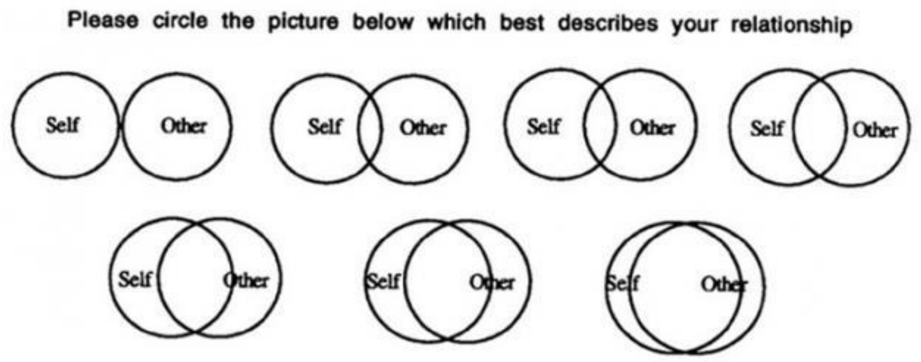

Figure 3. Inclusion of Other in Self measure.

Because Okel (2018) stated that the IOS scale was positively related to the Subjective Closeness Index (SCI), used in a study by Gächter, Starmer and Tufano (2015), this measure was used in this study as well. The SCI questionnaire consisted of two questions which were answered on a 7-point Likert scale where 1 means not close at all and 7 means very close.

The last part of the questionnaire consists of questions regarding emotion and the relationship between the participants in a dyad. The first question is generally whether the video triggers emotion, followed by an open question of what kind of emotion is triggered. Question regarding participant relationship asks how well do they know their experiment partner beforehand on a scale from 1 (i.e., *We never met before*) to 5 (i.e., *I know him/her very well*). The full questionnaire can be seen in Appendix C.

2.5 Procedure

Participants were recruited from the J.F. Schouten database. The selected participants were paired up and each of the pairs is randomly assigned to one of the two conditions. Before the experiment started, the participants were made to go through the informed consent and sign it. Afterwards, they were made to meet each other to make sure that they knew they are actually doing this experiment together with someone else. A brief explanation of the procedure was then given. They were informed that they will watch a relaxing video followed by an emotion-evoking video in the experiment. They were also informed that during the videos their heart-rate was being measured, and PS between them was being calculated during the latter video. Additionally, they were informed about the questionnaires they had to fill in at the end of the experimental video. Depending on the condition they were assigned to, further instructions regarding the colored border of the PS feedback and the intent of the colors were explicitly mentioned. The two experiment leaders then guided the participants to separate rooms to make sure they could not hear or see each other react to the emotion-evoking video to avoid influencing the results. Then, they received instructions on how to place the ECG electrodes of the Mobi device on their chest and assistance was provided upon request. When the participants had positioned the ECG electrodes and the experiment leaders had checked the signal, participants were instructed about pressing the marker button; to mark the start and the end of each video. Then, the participants put on the headphones which were used for sound output and the experimenters left the room. Shortly afterwards, the video was started at the same time for both participants. A schematic overview of the experiment procedure is visualized in Figure 4. Both the baseline video and the emotion-evoking video lasted for 4 minutes and 10 seconds. Instructions were given, within each video, about when to click the marker button; at the start and at the end of each video as explained before. After the baseline measurement was done, the emotion-evoking video started. Depending on the condition the dyad was assigned to, the level of PS was either shown or not (i.e., using the colored border). After the experimental video ended, the experiment leaders entered the rooms to provide the participants with the questionnaires and also to inform them that they could now remove the electrodes. The participants proceeded by filling in the questionnaires, which were provided on a laptop. After the questionnaires were filled in, the laptops were collected by the experimenters. Lastly, the participants were reunited, debriefed, compensated and thanked for their participation.

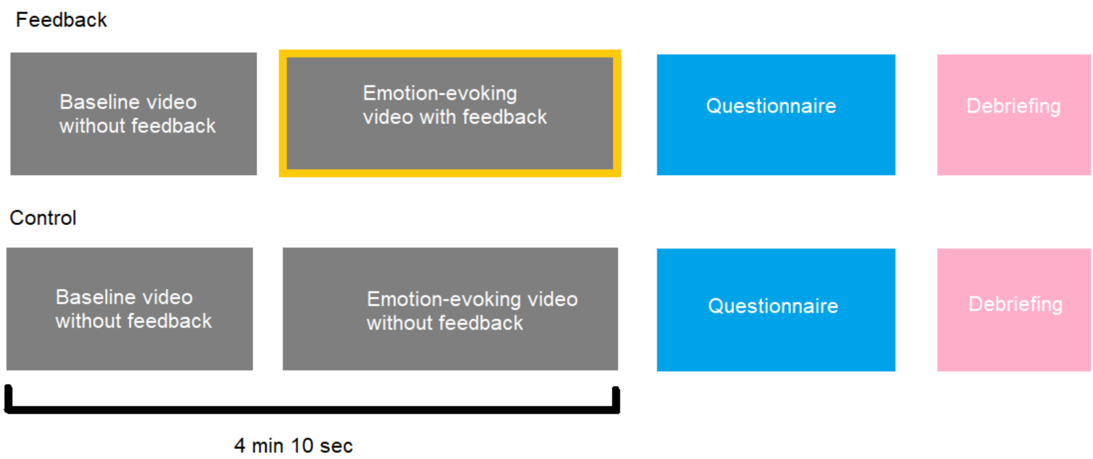

Figure 4. Schematic overview experiment procedure.

2.6 Pre-processing

After the experiment, data of each participant from all the sources (Mobi laptop and processors in the control room and participant room A) were linked together based on the participant IDs. The data consists of participant IDs, dyad numbers, demographic information, condition (feedback or no-feedback), empathy, IOS, SCI, and PS score. The empathy score was calculated by averaging the scores of the items in the questionnaire. The two measurements of social connectedness (IOS and SCI) were calculated separately, by averaging the scores of the corresponding items. These were done per person, in order to make differences in dyads visible.

For the PS, each participant was assigned a PS score, which is the same for the two participants in a dyad, that is the representative of their PS values during the emotion-evoking video. A temporal analysis of the PS values of all participants showed that PS values fluctuate drastically throughout the whole video session. A closer look at the median of PS values of feedback and no-feedback dyads across time revealed a timeframe of interest between 80 and 100 second of the emotion-evoking video (i.e., based on eyeballing the graph; see Figure 5). This timeframe had the least differences between the median PS scores of the feedback and the no-feedback dyads. This indicates the increasing and decreasing of arousal were in the same direction for both groups of dyads. Such homogeneity implied that all participants experienced something similar during this period, and therefore, this timeframe was specifically suitable to represent potential PS. The corresponding scene in the video showed the child being convinced by an older gentleman that his father is dead. The physiological reaction and scene together suggest that 80-100 seconds of the video is a more reliable emotion-evoking period, and the corresponding PS scores could be meaningful. Therefore, for the subsequent analyses, the median of sync values in this period was assigned as the PS score for each participant.

For the statistical analysis, only the results from dyads from which the participants do not know each other beforehand are reported. Dropping the participant dyads who knew each other, filters out potential external factors that could influence the relationship between PS, empathy, and PS feedback. A pre-existing connection between the participants in a dyad might result in biases. Therefore, data from participants who responded 2 or higher to the question about how well they know each other, meaning that they met the other participant before, were dropped to maintain a homogenous group. The analysis of the complete results, including participants that responded 2 or higher, can be found in Appendix D. Furthermore, all undesirable data (i.e., due to abnormal signals) and missing values (i.e., due to faulty questionnaire responses and participant dropout) found during the pre-processing phase were dropped as well, and the data belonging to those participants was not considered for the statistical analyses.

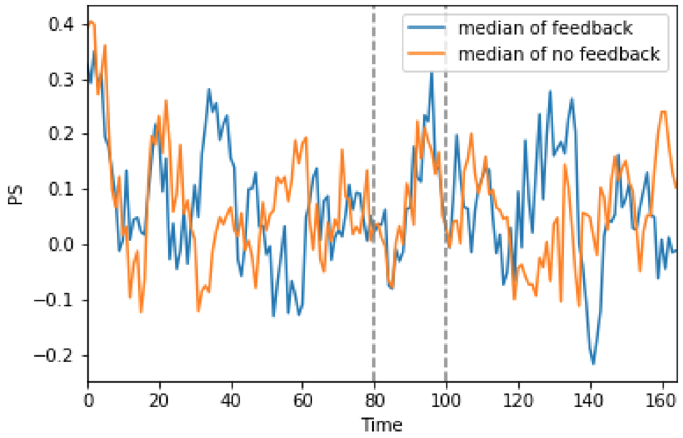

Figure 5. Median of sync values among participants across time.

2.7 Statistical analysis

For the first hypothesis, a Pearson Correlation test was conducted to investigate the relation between PS and perceived empathy as well as between PS and social connectedness (results can be found in the Appendix A for social connectedness). The tests were conducted on the synchronization data from the timeframe of interest, corresponding to the experimental video, irrespective of the condition the participant was assigned to. A post-hoc analysis was done to find correlation for the two conditions separately.

As for the second hypothesis, a moderated regression analysis was done with empathy as the dependent variable, PS value as the predictor variable and feedback as the moderator. The difference of PS-empathy slopes of the feedback and the no-feedback groups were understood as the coefficient of the interaction term of the moderated regression model. Post-hoc simple regression analyses were done for both the feedback and no-feedback groups separately to explain the effect of PS on empathy in isolation for the different conditions.

3. Results

For this study, 138 participants were measured, but only the data of 105 participants that did not know each other was used for analysis, due to drop out during the experiment, faulty measurements, and missing values (one participant did not fill in the questionnaire correctly). Of these 105 participants, 52 were male and 53 were female with an average age of 27.1 years (*SD* = 12.9). There were 56 and 49 participants in the no-feedback and feedback condition respectively. For the complete analyses including all the participants please refer to Appendix D.

As mentioned in the pre-processing section, the median of the PS values corresponding to a dyad and specific to the time frame of interest was considered as the PS score for both the participants in the dyad. The average PS score was .06 (*SD* = 0.26). Figure 6 displays a histogram of the PS score across both conditions. Also, the perceived empathy score per participant was calculated based on their responses in the questionnaire. Figure 7 displays a histogram of the empathy scores across both conditions. The average empathy score was 2.53 (*SD* = 0.41).

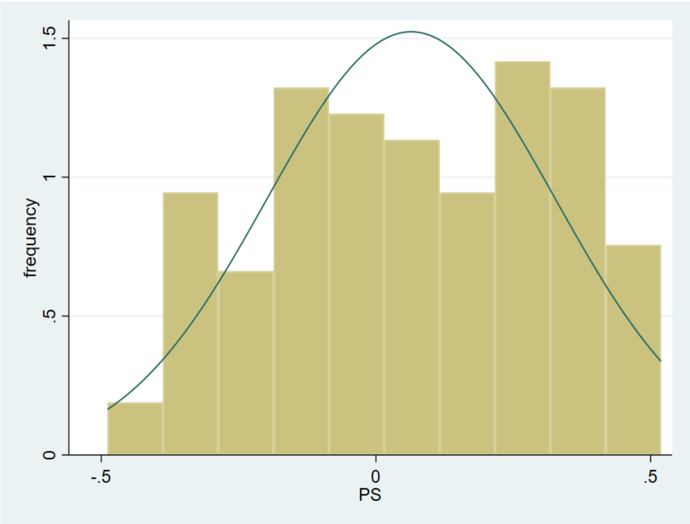

Figure 6. Histogram of PS score across both conditions.

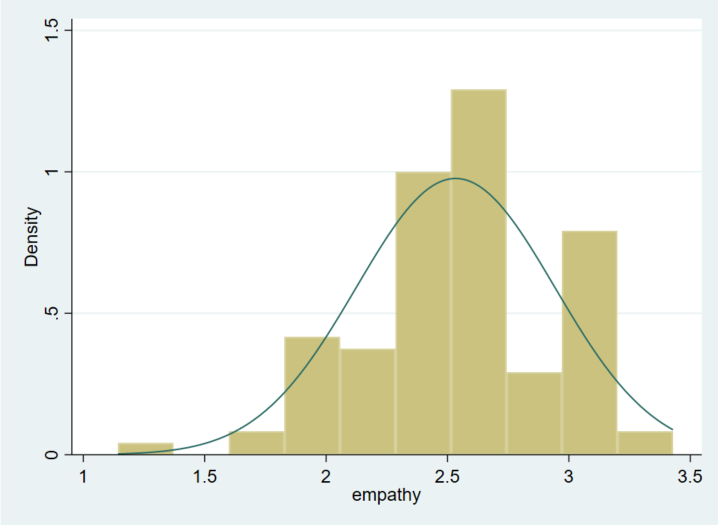

Figure 7. Histogram of empathy scores across both conditions.

Two separate analyses were carried out to test for both hypotheses. With regard to H1, a pairwise correlation measure was calculated. Results indicated that the correlation between PS score and empathy scores was not statistically significant ($r(103) = .07$, $p = .46$). Figure 8 visualizes this result as a scatter plot with an almost horizontal linear fit for the relation between empathy and PS scores, indicating no significant correlation.

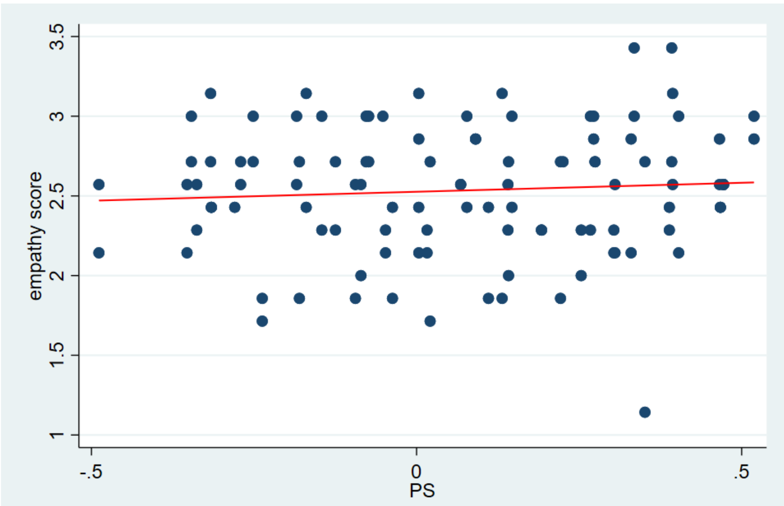

Figure 8. Scatter plot with a linear fit line for the relation between PS score and empathy for both the conditions combined.

While this was not a significant result, a post-hoc analysis involving the calculation of the pairwise correlation measure between PS score and empathy separately for the feedback group yielded a medium correlation with a significant result, ($r(47) = .41$, $p < .01$). Figure 9 visualizes this result as a scatterplot. Calculation of the pairwise correlation measure between PS and empathy for no-feedback condition did not yield a significant result, ($r(54) = -.11$, $p = .43$). This result is visualized as a scatterplot in Figure 10.

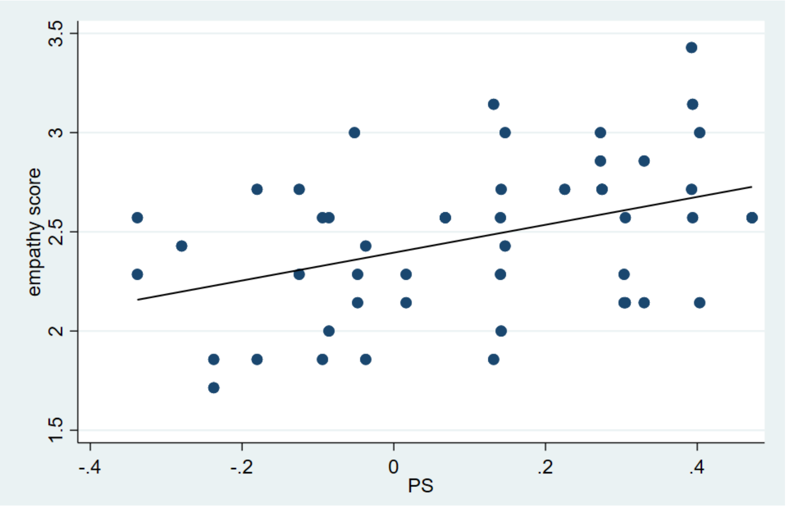

Figure 9. Scatter plot with a linear fit line for PS score and empathy for feedback condition.

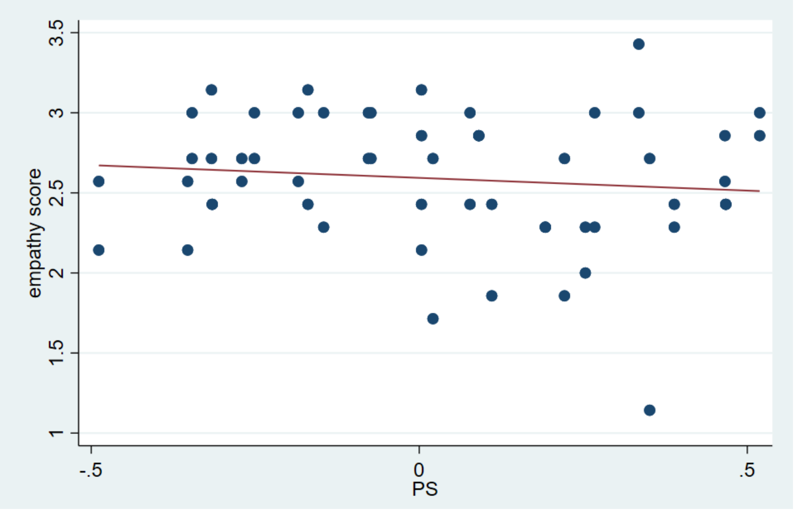

Figure 10. Scatter plot with a linear fit line for PS score and empathy for no-feedback condition.

With regard to H2, a moderated regression analysis with empathy as the dependent variable and PS scores as the predictor and the condition (feedback/no-feedback) as moderator was performed. The residuals had a normal distribution as determined with a

Shapiro-Wilk test, with $W = .99$ and $p = .30$. Both graphical inspection of the residuals, as well as a Breusch-Pagan / Cook-Weisberg test for heteroscedasticity, with $\chi 2(1) = 0.00$ and $p = .98$, suggested no violation of homoscedasticity. A few outliers were detected, both on the residuals as well as on the predictor. However, there were no influential data points, with all Cook's distances ≤ 0.14, and running the regression with outliers removed did not affect the interpretation of the regression outcomes. It was found that condition (feedback/no-feedback) and its interaction with PS scores predict empathy to a statistically significant extent, B_0 = 2.59 ($t(104) = 49.10$, p < .001), B_1(PS score) = -0.16 ($t(104) = -0.85$, $p = .40$), B_2(condition) = -0.20 ($t(104) = -2.44$, $p = .02$) and B_3(condition and PS interaction) = 0.86 ($t(104) = 2.77$, $p = .01$). B_3 reveals the difference in slopes of the linear model between the feedback and no-feedback condition. This difference in slopes between the two conditions is visualized in Figure 11 by means of a scatter plot and two separate regression lines corresponding to the two conditions. Feedback and its interaction with PS scores explained a significant proportion of variance in the empathy scores ($R^2 = .10$, $F(3, 101) = 3.71$, $p = .01$).

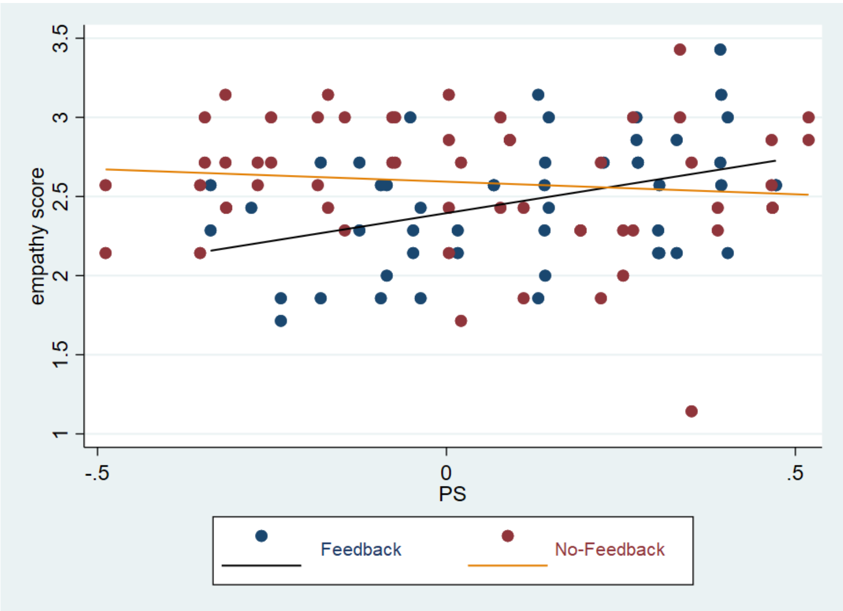

Figure 11. Scatter plot with a linear fit lines for PS and empathy for both no-feedback and feedback condition shown separately.

A post-hoc analysis of this linear model was conducted for the feedback and no-feedback conditions. This involved two separate simple regression analyses with PS scores as predictor and empathy as the dependent variable for both feedback and no-feedback conditions. In the feedback condition, the PS scores predict empathy to a statistically significant extent, B_0 = 2.40 (t(48) = 41.55, p < .001), and B_1(PS score) = 0.70 (t(48) = 3.04, p < .01). PS scores explained a significant proportion of variance in empathy scores (R^2 = .16, F(1, 47) = 9.25, p < .01). In the no-feedback condition, the PS scores do not predict empathy to a statistically significant extent, B_0 = 2.59 (t(55) = 46.40, p < .001), and B_1(PS score) = -0.16, (t(55) = -0.81, p = .42). PS scores did not explain a significant proportion of variance in empathy scores (R^2 = .01, F(1, 54) = 0.65, p = .42).

4. Discussion

4.1 General discussion

A large portion of communication nowadays is done remotely through text, image, and video messaging. Since most non-verbal cues cannot be delivered, people cannot convey themselves or perceive others as much as they do in face-to-face interaction, since perception of empathy relies heavily on these cues (Montague et al., 2013). To address this issue, the current research focused on synchrony of humans' physiological states, since it is claimed to reflect empathy. An investigation was done about how presenting the level of mutual synchrony would influence people's perception of each other.

For the current study, the goal was to determine whether PS has a positive relationship with perceived empathy (H1), and whether this relationship has a different slope for the feedback condition in comparison with the no-feedback condition (H2). Only the data of participants who did not know each other were used for the analysis to ensure homogeneity in participants. When looking at the relationship between the PS scores across both conditions and perceived empathy, no significant overall relationship was found. H1 can therefore be rejected. This is not in line with previous other studies (Marci et al., 2007; Messina et al., 2013). These studies did indeed find overall significant relationships between PS and perceived empathy. However, in these experiments, PS was calculated by means of skin conductance. Due to time constraints and practical reasoning, the current study chose to employ synchronization in HR variability as a measure of PS. This could be a potential reason for the deviation of results from this study compared to prior literature. The study on HR verifies that not all types of physiological measurements result in the same effect. Further, another explanation of the inconsistent outcomes could be that participants were subjected to different tasks in the experiments. Also, these studies involved tasks that had prolonged interaction between the participants, whereas the current study did not.

A closer look at PS and empathy shows that, when the pairwise correlation analysis was performed on two conditions separately, in the feedback condition the positive correlation between PS and perceived empathy was actually significant. This is somewhat in line with Levenson and Gottman (1983), as they investigated PS influences during the day in married couples. They found nonsignificant results for effects of PS on their satisfaction of marriage during different events of the day. In contrast, they found significant effects of PS on perceived satisfaction during problematic periods of time. According to Levenson and Gottman (1983), effects of PS on perception can be different during different moments. The same might go for the results of the current study, as receiving feedback or not is a different experience. Therefore, even though H1 is rejected, it is important to note that the PS and perceived empathy are positively related in the case of PS feedback being provided.

A possible explanation for the inconsistent results in two conditions is that the participants did not really talk to each other before the experiment started. This did not pose any problem for the feedback condition as the participants might have felt more connected to each other. It could be that the participants in the feedback condition considered the feedback to be a nonverbal cue. On the contrary, in the no-feedback condition, participants asked questions of the following nature while filling in the questionnaire, “What do I need to fill in? I don’t know this!”. Since they did not receive any information of the person they were participating in the experiment with, it could be inferred that this made it harder for them to fill in questions about the empathy they have towards each other. Therefore, this could explain the significant result between PS and empathy score in the feedback condition.

A moderated regression analysis was performed to find out the contribution of feedback to the relationship between PS and empathy. The main effect of feedback and the interaction effect of feedback and PS scores were found to be significant. The main effect of PS feedback indicates that feedback is negatively associated with perceived empathy. However, the negative coefficient could be due to the unequal group membership between high and low sync dyads which could not be controlled. This was because the main criterion for the feedback thresholds was not to balance the number of high and low sync dyads but to ensure that the high sync groups saw a 60-30 green-red ratio in their feedback and vice versa for the low-sync groups. The significant interaction effect shows that the slope of the regression line of the feedback condition is different from (and steeper than) the slope of the no-feedback condition. Hence, H2 can be accepted.

Further, post-hoc analyses involving separate simple regression analyses for the feedback and no-feedback condition with empathy as the dependent variable and PS score predictor helped understand the influence of PS scores in a clearer fashion. In the feedback condition, the coefficient of PS score was found to be significant and positive. From this, it could be inferred that when PS level was presented as feedback, it was treated as a nonverbal cue and provided some information for the participants to form an opinion of their experiment partners.

4.2 Implications

The findings in this study can help in understanding and utilizing the influence of PS in remote communications. The results provide insights on the relationship between PS and empathy. The positive correlation between the two variables while PS feedback is presented suggests that PS could be an indicator on empathy, as higher PS is associated with a higher level of synchrony especially when feedback is presented. PS as a quantifiable parameter is easier to transmit compared to the nonverbal cues. Thus, it could be a possible tool for raising empathy and thus improving computer-mediated communications if future research proves this positive relation to be causal. Besides, the role of PS feedback is crucial, since the results show that PS-empathy relationship is significant in the feedback condition. Moreover, the significant moderated effect of PS feedback proves to be an influential factor that could alter such relationship. In the future, if the scientific community is convinced that presenting PS leads to a higher level of empathy, sharing such information in online communication can bring people closer to each other. For example, as empathy is a significant predictor of treatment results in psychotherapy, increasing empathy through PS feedback could benefit therapy outcomes and medical relations (Mohr, 1995). Other applications in contexts such as online dating, client meetings, and remote team building in companies are foreseeable.

4.3 Limitations

Several limitations regarding the procedure of the experiment surfaced while analyzing the data. As mentioned earlier, though the participants met each other before the experiment, their interaction was quite limited in the manner that they did not develop any familiarity. Because the experimental partners hardly interacted, they had no basis to assess the perceived empathy when they were asked to do so. Also, two dyads in the feedback condition mentioned that they could have been in low-sync during the debriefing, because they remember the border to be more red than green or yellow. After the analysis, it was found that these dyads were actually in the high-sync group. This indicates that participants could be biased towards the overall experience, in other words, what they felt is not necessarily the same as what was presented. To control this bias, a question regarding the perceived overall level of PS (i.e., high or low) could be included in the questionnaire, and participants whose perceived level is not in line with actual data could be identified.

Furthermore, it had been brought to notice that two dyads believed that the feedback was simulated which was later clarified that it was indeed calculated in real-time during debrief. Including a measure of how the participants perceive their synchrony could give more insights into how feedback is perceived. This can also be used to account for the above mentioned problem of how the dyads perceive their level of synchrony. As a control variable, the perceived level of synchronization could help in constructing a better model in understanding the effect of feedback on perceived empathy.

Considering only two studies (Marci et al., 2007; Messina et al., 2013) to determine the sample size for the current study, may have led to the study being underpowered. Only a few studies investigated the relationship between PS and empathy, and in both those studies, skin conductance was used as a measure of PS.

Another limitation is that the video used in this experiment (The Champ, 1979) might not be emotional to the younger participants. Based on the responses from the questionnaire, 28% of the participants reported that they did not feel emotional after watching the emotion-evoking video. Further, it should be noted that among those participants, 93% were younger than 30 years leading to the inference that the video may not be emotional for the younger demographic. Also, two participants reported they laughed while watching the video. The goal of this video was to evoke a common emotion in the participating dyad and a sadness evoking video was chosen in particular as affectively unpleasant stimuli tend to generate greater arousal (Hazer et al., 2015; Hewig et al., 2005, Gross & Levenson, 1995). Since the average age of participants in this study was 27.1 years and the video was filmed in 1979, most participants are habituated to videos with a different style of acting, which may be the reason why this video was not as emotion-evoking as expected. For future research, a different emotion-evoking video should be considered, depending on the target group of the study.

Having participants perform discussion tasks before watching the video could solve the problem of participants not knowing what to fill in the question regarding familiarity with each other in the questionnaire. Participants get to know each other more by having discussions. However, this discussion should be led by the experimenter to control the subjects of discussion. A pre and post questionnaire to measure the change in closeness can also give more insights in how participants' perceived relationship can change during the experiment. Further analysis could be done to understand whether familiarity has influence on the PS-empathy relationship.

Additionally, follow-up studies can consider a different strategy for determining the thresholds that dictate high and low synchrony. In the current study a ratio of 60-30 for green-red was chosen for high sync groups and vice versa, relying on previous work of Van Laar (2019) and the pilot study. This threshold selection criteria results in having only 20 high sync dyads among 105 dyads. The large inequality in the group sizes pertaining to level of synchrony could bias the statistical analysis. Future researchers should consider reducing this inequality in size between the two groups, while keeping in mind that the percentages of each color displayed distinctly convey the overall level of synchrony between the dyad.

5. Conclusion

The current study investigated whether there is a relationship between PS and perceived empathy and how feedback can be a moderator of this effect. The experiment that was conducted consisted of dyads of participants that were randomly divided into one of two conditions: feedback and no-feedback condition. Participants watched an emotion-evoking video, while their PS values were being measured. The feedback condition received feedback about how high or low they were in synchrony by means of a colored border around the video. A questionnaire reported their empathy towards each other and questions about the variable social connectedness. Results conclude that there is a positive relationship between PS and perceived empathy, but only within the feedback condition. Moderated regression analysis showed that feedback is a significant moderator in this relationship between PS and perceived empathy. This study shed a light on PS, but more in depth research should be performed, predominantly with a bigger sample size, to gain more knowledge about the relationship of PS and empathy and the effects of feedback. It is recommended that participants meet before the experiment and preferably carry out a task together to ensure a slightly prolonged interaction to establish a bond of familiarity between them.

6. References

Appelhans, B., & Luecken, L. 2006. Heart Rate Variability as an Index of Regulated Emotional Responding. Review of General Psychology, 10(3), 229-240. doi: 10.1037/1089-2680.10.3.229

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology, 63*(4), 596-612. doi:10.1037/0022-3514.63.4.596

Batson, C. D., Duncan, B. D., Ackerman, P., Buckley, T., & Birch, K. (1981). Is empathic emotion a source of altruistic motivation? *Journal of Personality and Social Psychology*, *40*(2), 290–302. doi:10.1037/0022-3514.40.2.290

Boschman, M. C. (2017). Acquisition of Physiological Data @ HTI - User Manual. Eindhoven: Technische Universiteit Eindhoven Human Technology Interaction Group.

Boschman, M. C. (2018). HRSync User Manual. Eindhoven: Technische Universiteit Eindhoven Human Technology Interaction Group.

Carter, K. A. (2003). Type me how you feel: Quasi-nonverbal cues in computer-mediated communication. *et Cetera*, *60*(1), 29-39.

Cramer, D., & Jowett, S. (2010). Perceived empathy, accurate empathy and relationship satisfaction in heterosexual couples. *Journal of Social and Personal Relationships*, *27*(3), 327–349. doi:10.1177/0265407509348384

Cuff, B. M., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: a review of the concept. *Emotion Review*, *8*(2), 144-153. doi:10.1177/1754073914558466

DiMascio, A., Boyd, R. W., & Greenblatt, M. (1957). Physiological Correlates of tension and Antagonism during Psychotherapy: A study of Interpersonal Physiology. *Psychosomatic medicine*, *19*(2), 99-104. doi:10.1097/00006842-195703000-00002

Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the closeness of relationships: a comprehensive evaluation of the Inclusion of the Other in the Self scale. *PloS One, 10*(6). doi:10.1371/journal.pone.0129478

Hatfield, E., Rapson, R. L., Ye, Y. C., Decety, J., & Ickes, W. (2009). *The social neuroscience of empathy*. Boston: MIT Press. doi:10.7551/mitpress/9780262012973.001.0001

Hazer, D., Ma, X., Rukavina, S., Gruss, S., Walter, S., & Traue, H. C. (2015, August). Emotion elicitation using film clips: Effect of age groups on movie choice and emotion rating. In *International Conference on Human-Computer Interaction* (pp. 110-116). Springer, Cham. doi:10.1007/978-3-319-21380-4_20

Hewig, J., Hagemann, D., Seifert, J., Gollwitzer, M., Naumann, E., & Bartussek, D. (2005). A revised film set for the induction of basic emotions. *Cognition and emotion*, *19*(7), 1095. doi::10.1080/02699930541000084

Hoffman, M. L. (1984). Interaction of affect and cognition in empathy. In C. E. Izard & R. B. Kagan (Eds.), *Emotions, cognition, and behavior* (pp. 103-131). Cambridge: Cambridge University Press.

Janssen, J. H., Bailenson, J. N., IJsselsteijn, W. A., & Westerink, J. H. (2010). Intimate heartbeats: Opportunities for affective communication technology. *IEEE Transactions on Affective Computing*, *1*(2), 72-80. doi:10.1109/T-AFFC.2010.13

Levenson, R. W., & Gottman, J. M. (1983). Marital interaction: physiological linkage and affective exchange. *Journal of personality and social psychology, 45*(3), 587-597. doi:10.1037/0022-3514.45.3.587

Marci, C. D., Ham, J., Moran, E., & Orr, S. P. (2007). Physiologic correlates of perceived therapist empathy and social-emotional process during psychotherapy. *The Journal of nervous and mental disease*, *195*(2), 103-111. doi:10.1097/01.nmd.0000253731.71025.fc

Ménard, M., Richard, P., Hamdi, H., Daucé, B., & Yamaguchi, T. 2015. Emotion Recognition Based on Heart Rate and Skin Conductance. Proceedings of the 2nd International Conference on Physiological Computing Systems,1, 26-32. doi:10.5220/0005241100260032

Messina, I., Palmieri, A., Sambin, M., Kleinbub, J. R., Voci, A., & Calvo, V. (2013). Somatic underpinnings of perceived empathy: The importance of psychotherapy training, *Psychotherapy Research*, *23*(2), 169-177. doi:10.1080/10503307.2012.748940

Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, *103*(3), 324-344. doi:10.1037/0033-2909.103.3.324

Mohr, D. C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical psychology: Science and practice*, *2*(1), 1-27. doi:10.1111/j.1468-2850.1995.tb00022.x

Montague, E., Chen, P. Y., Xu, J., Chewning, B., & Barrett, B. (2013). Nonverbal interpersonal interactions in clinical encounters and patient perceptions of empathy. *Journal of Participatory Medicine*, 5. Retrieved from: https://participatorymedicine.org/journal/evidence/2013/08/14/nonverbal-interpersonal-interactions-in-clinical-encounters-and-patient-perceptions-of-empathy/

Okel, S. E. (2018). *The Effect of Biofeedback on Perceived Empathy and Social Connectedness during social interaction.* Eindhoven: Eindhoven University of Technology.

Palumbo, R. V., Marraccini, M. E., Weyandt, L. L., Wilder-Smith, O., McGee, H. A., Kiu, S., & Goodwin, M. S. (2016). Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, *21*(2), 99-141. doi:10.1177/1088868316628405

Preece, J. (1999). Empathy online. *Virtual Reality*, *4*(1), 74-84. doi:10.1007/BF01434996

Preston, S. D., & De Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and brain sciences*, *25*(1), 1-20. doi:10.1017/S0140525X02000018

Plank, R. E., Minton, A. P., & Reid, D. A. (1996). A short measure of perceived empathy. *Psychological reports*, *79*(3), 1219-1226. doi:10.2466/pr0.1996.79.3f.1219

Smith, D. B. D., & Strawbridge, P. J. 1968. Stimulus duration and the human heart rate response. Psychonomic Science, 10(2), 71-72. Retrieved from: doi:10.3758/BF03331412

Van Bel, D. T., IJsselsteijn, W. A., & Kort, de, Y. A. W. (2008). Interpersonal connectedness: conceptualization and directions for a measurement instrument. *Proceedings of the 26th CHI Conference on Human Factors in Computing Systems*, April 5-10, 2008, Florence, Italy (pp. 3129-3134). New York: Association for Computing Machinery, Inc. doi: 0.1145/1358628.1358819

Van Laar, K. C. K. (2019). *The moderating effect of Biofeedback on the relation between Physiological Synchrony and Empathy & Team Performance.* Eindhoven: Eindhoven University of Technology.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological science, 20*(1), 1-5. doi:10.1111/j.1467-9280.2008.02253.x

7. Appendix

Appendix A. Statistical analysis of Social connectedness[1]

Since SCI and IOS both measure the same construct, i.e., Social connectedness, an average of both SCI and IOS measures was considered as Social Connectedness score. The mean value of the Social connectedness score was 2.02 (*SD* = 1.04). Figure 12 displays a histogram of the Social connectedness score across both conditions.

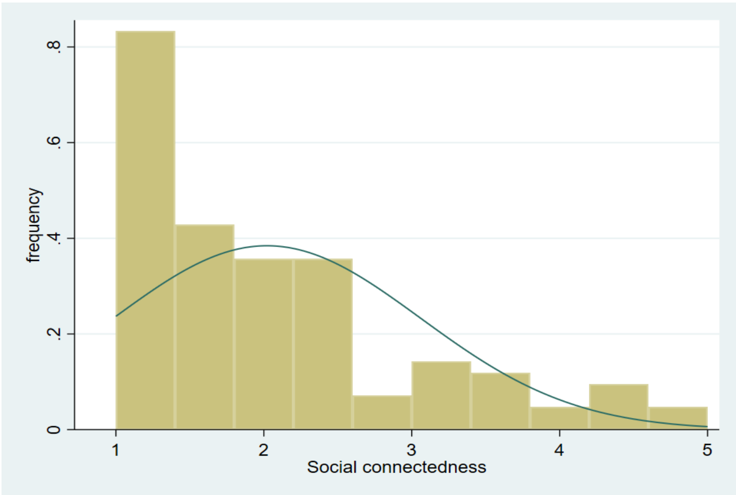

Figure 12. Histogram of Social Connectedness.

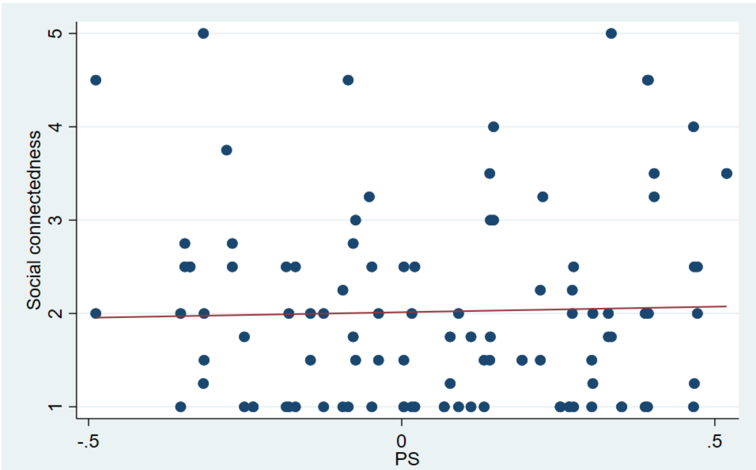

Figure 13. Scatter plot with a linear fit line for PS score and Social connectedness for both the conditions.

[1] only for participants who are not familiar with each other

A pairwise correlation measure was calculated. Results indicated that the correlation between PS score and Social connectedness score was not statistically significant ($r(103) = .03$, $p = .76$). Figure 13 visualizes this result as a scatter plot with an almost horizontal linear fit for the relation between Social connectedness and PS scores, indicating no significant correlation. Further, separate post-hoc analyses for both feedback and no-feedback conditions did not reveal any significant results. For the feedback condition, the pairwise correlation measure between PS score and Social connectedness resulted in $r(47) = .17$ ($p = .243$). Figure 14 visualizes this result as a scatterplot. Calculation of the pairwise correlation measure between PS and Social Connectedness for no-feedback condition did not yield a significant result, $r(54) = -.09$ ($p = .498$). This result is visualized as a scatterplot in Figure 15.

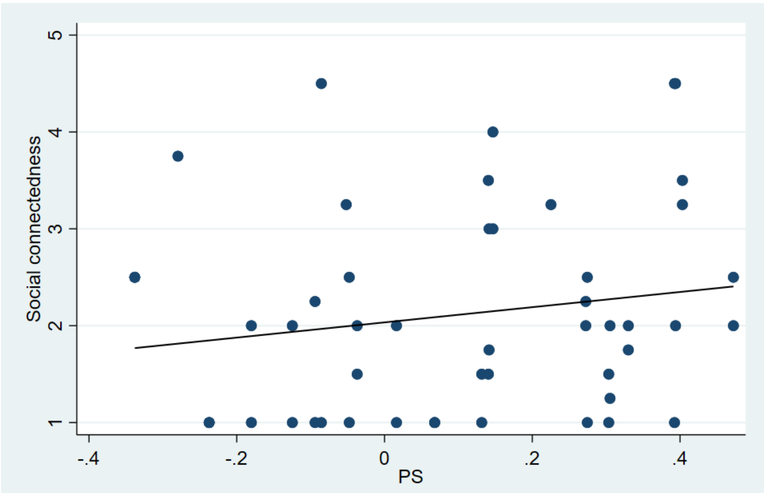

Figure 14. Scatter plot with a linear fit line for PS score and Social connectedness for feedback condition.

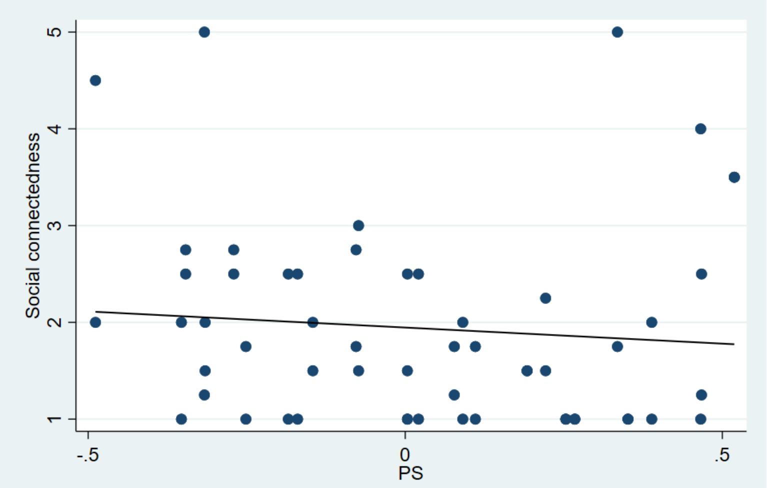

Figure 15. Scatter plot with a linear fit line for PS score and Social connectedness for no-feedback condition.

A moderated regression analysis with Social connectedness as the dependent variable and PS scores as the predictor and the condition (feedback/no-feedback) as moderator was performed. The residuals did not have a normal distribution as determined with a Shapiro-Wilk test, with $W = .88$ and $p < .001$. Both graphical inspection of the residuals, as well as a Breusch-Pagan / Cook-Weisberg test for heteroscedasticity, with $\chi 2(1) = 0.02$ and $p = .901$, suggested no violation of homoscedasticity. A few outliers were detected, both on the residuals as well as on the predictor. However, there were no influential data points, with all Cook's distances ≤ 0.33, and running the regression with outliers removed did not affect the interpretation of the regression outcomes. Due to the violation of normality assumption, a robust moderated regression was performed. It was found that condition (feedback/no-feedback) and PS scores do not predict Social connectedness to a statistically significant extent, $B_0 = 1.95$ ($t(104) = 13.95$, p < .001), B_1(PS score) = -0.33 ($t(104) = -0.68$, $p = .5$), B_2(condition) = 0.09 ($t(104) = 0.41$, $p = .685$) and B_3(condition and PS interaction) = 1.12 ($t(104) = 1.36$, $p = .176$). B_3 indicates the difference in slopes of the linear model between the feedback and no-feedback condition. This difference in slopes between the two conditions is visualized in Figure 16 by means of a scatter plot and two separate regression lines corresponding to the two conditions. Feedback and PS scores do not explain a significant proportion of variance in the Social connectedness scores (R^2 = .03, F(3, 101) = 0.89, p = .449).

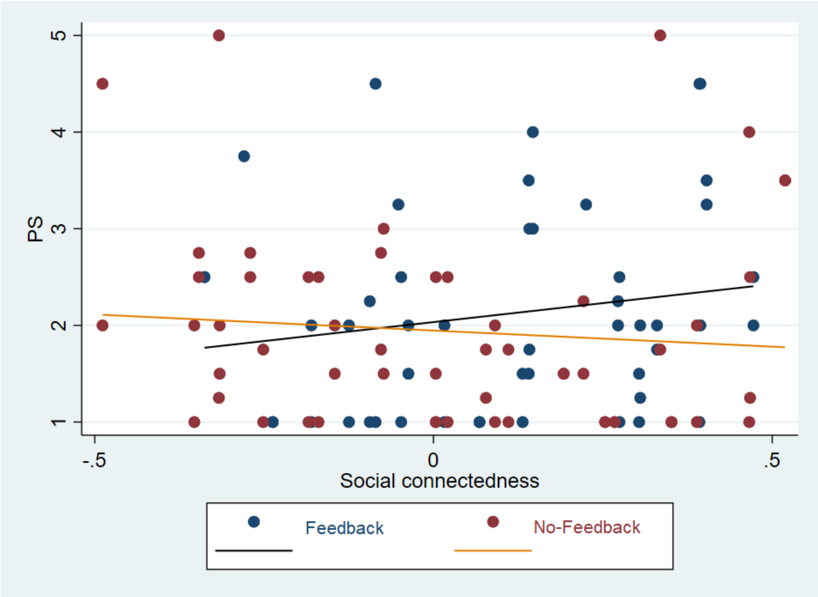

Figure 16. Scatter plot with a linear fit lines for PS and Social connectedness for both no-feedback and feedback condition shown separately.

A post-hoc analysis of this linear model was conducted for the feedback and no-feedback conditions. This involved two separate simple regression analyses with PS scores as predictor and Social connectedness as the dependent variable for both feedback and no-feedback conditions. In the feedback condition, the PS scores do not predict Social connectedness to a statistically significant extent, B_0 = 2.03 (t(48) = 12.28, p < .001), and B_1(PS score) = 0.78 (t(48) = 1.18, p = .243). PS scores did not explain a significant proportion of variance in Social connectedness scores (R^2 = .02, F(1, 47) = 1.40 , p = .243). In the no-feedback condition, the PS scores do not predict empathy to a statistically significant extent, B_0 = 1.94 (t(55) = 14.08, p < .001), and B_1(PS score) = -0.33, (t(55) = -0.68, p = .498). PS scores did not explain a significant proportion of variance in Social connectedness scores (R^2 = .01, F(1, 54) = 0.47, p = .497).

Appendix B. *G*Power* sample size calculation

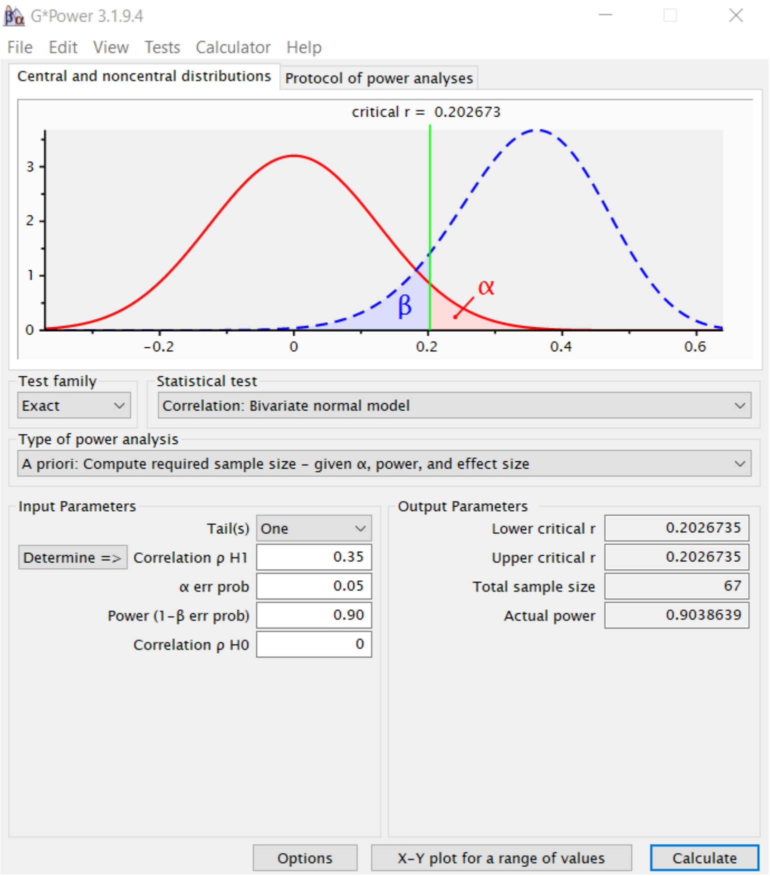

Figure 17. Screenshot of G*Power sample size calculation.

Appendix C. Questionnaires

Empathy measure used by Okel (2018) & Van Laar (2019) based on Plank et al. (1996)

1. This persons understands me and my role in this experiment.
2. I have lousy feelings when dealing with this person.
3. This person really understood my feelings about this situation.
4. I feel as if I am on the same wavelength as this person.
5. This person does not understand how I think.
6. This person has a lot of knowledge about how I need to make decisions.
7. This person always understood my needs.

SCI measure used by Okel (2018) based on Gächter and colleagues (2015)

1. Relative to all your other relationships (both same and opposite sex) how would you characterize your relationship with the other participant?
2. Relative to what you know about other people's close relationships, how would you characterize your relationship with the other participant?

Other questions

1. Did the video trigger your emotions?
2. If yes, what kind of emotions did the video trigger?
3. How well did you know the other participant before the experiment?

Appendix D. Statistical analysis on perceived empathy for all participants

This analysis includes participants who were familiar with each other. The mean value of the PS score was 0.06 (*SD* = 0.26). Figure 18 displays a histogram of the PS score across both conditions. Also, the perceived empathy score per participant was calculated based on their responses in the questionnaire. Figure 19 displays a histogram of the empathy scores across both conditions. The mean value of the empathy was 2.60 (*SD* = 0.44).

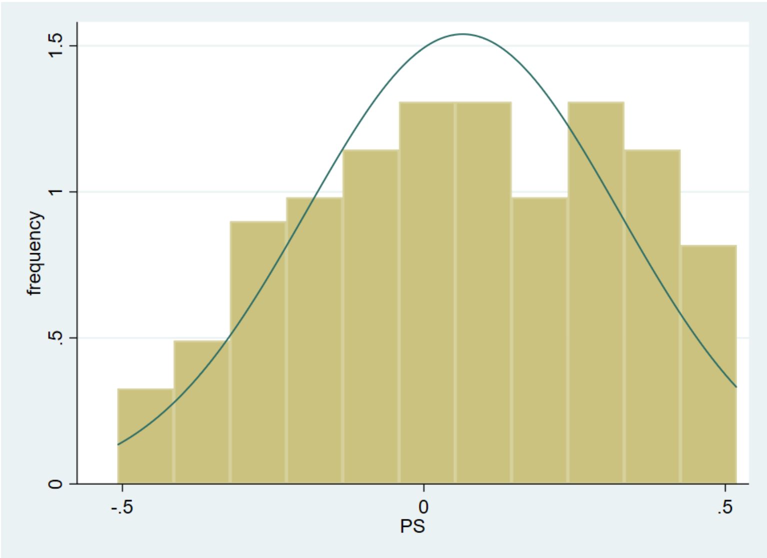

Figure 18. Histogram of PS across both conditions.

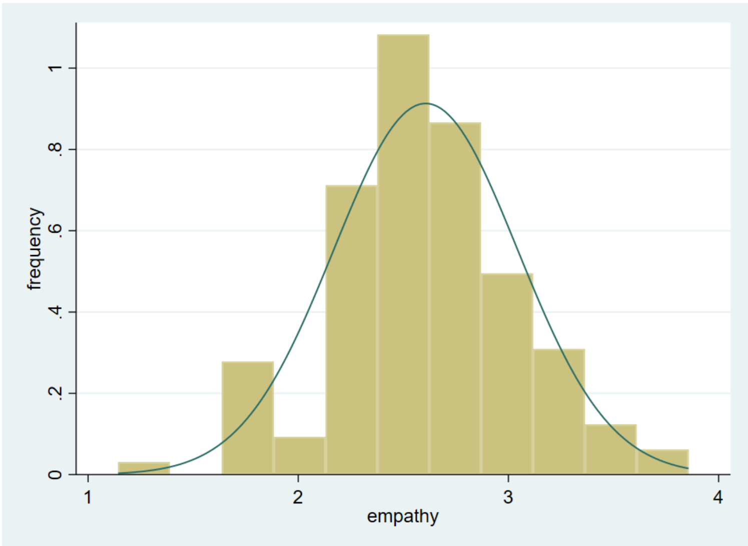

Figure 19. Histogram of empathy scores across both conditions.

Two separate analyses were carried out to test for both hypotheses. With regard to H1, a pairwise correlation measure was calculated to measure the correlation between PS score and empathy scores, *r*(129) = .09 (*p* = .319). Figure 20 visualizes this result as a scatter plot with an almost horizontal linear fit for the relation between empathy and PS scores, indicating no significant correlation.

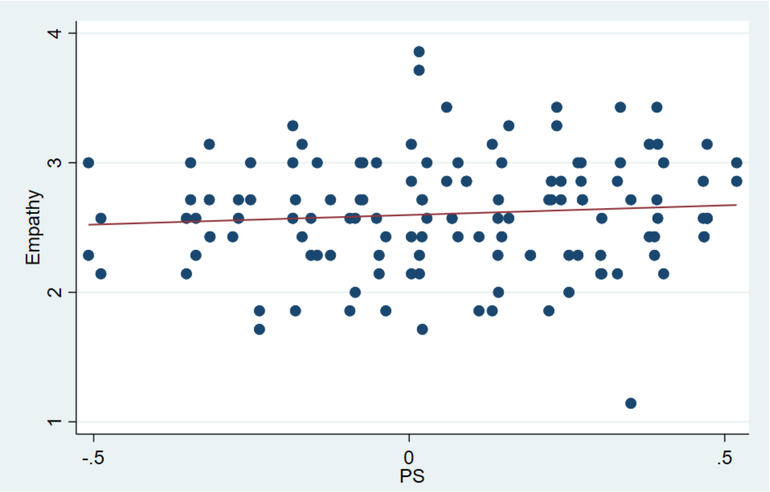

Figure 20. Scatter plot with a linear fit line for PS score and empathy for both the conditions.

While this did not yield a significant result, a post-hoc analysis involving the calculation of the pairwise correlation measure between PS score and empathy separately for the feedback group yielded a significant result, *r*(63) = .28 (*p* = .025). Figure 21 visualizes this result as a scatterplot. Calculation of the pairwise correlation measure between PS and empathy for no-feedback condition did not yield a significant result *r*(64) = -.06 (*p* = .656). This result is visualized as a scatterplot in Figure 22.

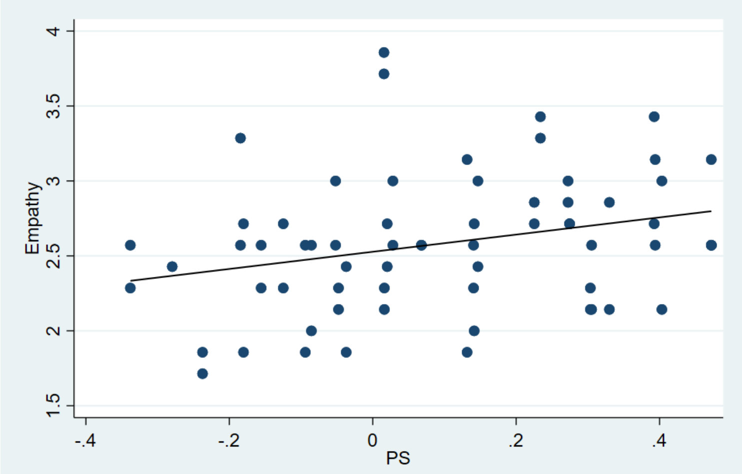

Figure 21. Scatter plot with a linear fit line for PS score and empathy for feedback condition.

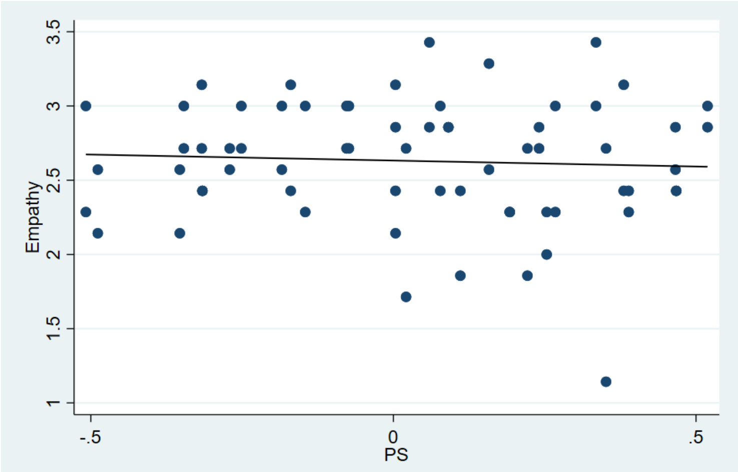

Figure 22. Scatter plot with a linear fit line for PS score and empathy for no-feedback condition.

With regard to H2, a moderated regression analysis with empathy as the dependent variable and PS scores as the predictor and the condition (feedback/no-feedback) as moderator was performed. The residuals had a normal distribution as determined with a Shapiro-Wilk test, with $W = .99$ and $p = .432$. Both graphical inspection of the residuals, as well as a Breusch-Pagan / Cook-Weisberg test for heteroscedasticity, with $\chi 2(1) = 0.82$ and $p = .365$, suggested no violation of homoscedasticity. A few outliers were detected, both on the residuals as well as on the predictor. However, there were no influential data points, with all Cook's distances ≤ 0.10, and running the regression with outliers removed did not affect the interpretation of the regression outcomes. It was found that the interaction between PS scores and condition predicts empathy to a statistically significant extent, B_0 = 2.63 (t(130) = 49.21, p < .001), B_1(PS score) = -0.08 (t(130) = -0.44, p = .664), B_2(condition) = -0.10, (t(130) = -1.33, p = .187), and B_3(condition and PS interaction) = 0.65 (t(130) = 2.14, p = .034). B_3 reveals the difference in slopes of the linear model between the feedback and no-feedback condition. This difference in slopes between the two conditions is visualized in Figure 23 by means of a scatter plot and two separate regression lines corresponding to the two conditions. Only the interaction of condition with PS scores explained a significant proportion of variance in the empathy scores (R^2 = .05, F(3, 127) = 2.07, p = .108).

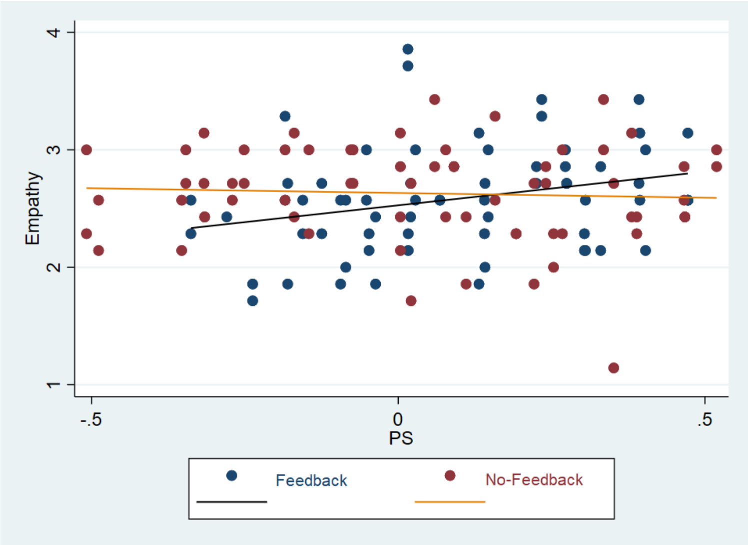

Figure 23. Scatter plot with a linear fit lines for PS and empathy for both no-feedback and feedback condition shown separately.

A post-hoc analysis of this linear model was conducted for the feedback and no-feedback conditions. This involved two separate simple regression analyses with PS scores as predictor and empathy as the dependent variable for both feedback and no-feedback conditions. In the feedback condition, the PS scores predict empathy to a statistically significant extent, B_0 = 2.53 (t(64) = 42.17, p < .001), and B_1(PS score) = 0.57 (t(64) = 2.31, p = .024). PS scores explained a significant proportion of variance in empathy scores (R^2 = .08,

$F(1, 63) = 5.31$ $p = .024$). In the no-feedback condition, the PS scores do not predict empathy to a statistically significant extent, $B_0 = 2.63$ ($t(65) = 50.62$, $p < .001$), and B_1(PS score) = -0.08 ($t(65) = -0.45$, $p = .656$). PS scores did not explain a significant proportion of variance in empathy scores ($R^2 = .00$, $F(1, 64) = 0.20$, $p = .656$).