A Talking Musical Robot during Long-term Interaction

After Bonding and Empathy Fade, Relevance and Realism Arise

(Tech. Rep.)

Ivy S. Huang

Laboratory for Artificial Intelligence in Design

The Hong Kong Polytechnic University

Introduction

Based on literature review, we assumed musical robots may be an alternative to talking robots to deliver empathy, given that current NLP is limited and (popular) music is appealing to youth. This assumption brought three research questions: 1) To what extent can individuals perceive artificial empathy when interacting with on-screen robots? 2) Which are the mechanisms involved in conveying robot empathy to its users? 3) How does the experience of interacting with an empathic robot change at different timepoints? 4) Over time, what role does music play in conveying robot empathy? This report aimed to describe the experiment that queried the hypotheses and offer the elaboration of the statistical analyses of results.

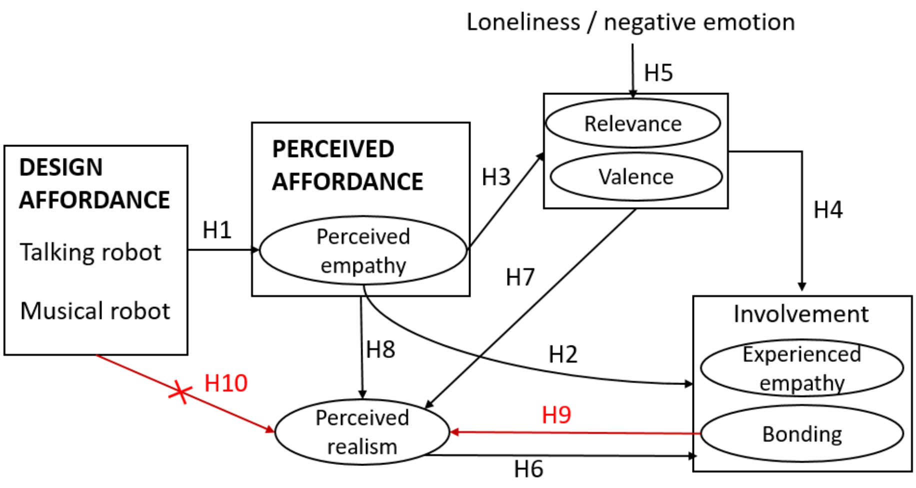


Figure 1. *I-PEFiC and TAB combined for an empathic robot*

-

Our work started from the frameworks called I-PEFiC (Van Vugt, Hoorn, & Konijn, 2009) and TAB (Konijn & Hoorn, 2017), describing how showing empathy through music or talking may influence a participant's perception and experience of a robot. The syntheses of these theories can be found in Table 1.

Table 1

Summary of Research Questions and Hypotheses

Research Questions	Hypotheses

RQ1: *To what extent can individuals perceive artificial empathy when interacting with on-screen robots?*	H1: Humans can perceive empathy expressed through music and speech by on-screen robots.
	H1a: Combining music and speech results in greater perceived empathy than either component alone.
	H1b: When compared to robots that express empathy through music only, robots expressing empathy verbally as well as through music elicit a stronger perception of empathy.
RQ2: *Which are the mechanisms involved in conveying robot empathy to its users?*	H2: Relevance and outcome expectations (valence) mediate the relationship between perceived empathy and involvement.
	H2a: Greater perceived empathy leads to stronger experiences of reciprocated empathy.
	H2b: Greater perceived empathy leads to stronger experiences of bonding.
	H3a: The robot's relevance to personal concerns increases with perceived empathy.
	H3b: Positive expectations (valence) increase with perceived empathy.
	H4a: More relevance elicits stronger involvement (including reciprocated empathy and bonding).
	H4b: Higher expectations evoke stronger involvement (including reciprocated empathy and bonding).
	H5: The impact of perceived empathy is amplified by the individual's level of loneliness.
	H6: Perceived realism positively contributes to involvement (including reciprocated empathy and bonding).
	H7: Perceived realism is positively influenced by perceived empathy.
	H8: Perceived realism is positively influenced by expectation valence.
	H9: Stronger involvement (including reciprocated empathy and bonding) feeds back into higher perceived realism.
	H10: By design, a robot's specific actions (like playing a trumpet) do not have a direct effect on its perceived realism.
RQ3: *How does the experience of interacting with an empathic robot change at different timepoints?*	H11a: Perceived empathy may decrease over time.
	H11b: Positive expectations (valence) may decrease over time.
	H11c: An empathic robot stays relevant to personal concerns.
	H12a: Reciprocated empathy persists over time.
	H12b: Bonding tendencies persist over time.
RQ4: *Over time, what role does music play in conveying robot empathy?*	H13: Musical empathy may be less noticeable initially but becomes more apparent over time.
	H14a: Over time, music exerts more reciprocated empathy than conversation
	H14b: Over time, music exerts stronger bonding than conversation
	H14c: Music evokes greater perceived realism and contributes to bonding, enabling users to recognize empathy in the robot during subsequent interactions.

Method

Participants and Design

We conducted an experiment with a 2 (talk: yes vs. no) × 2 (music: yes vs. no) repeated (3 times) factorial design to understand the participants' perception and experience of a robot while it expresses empathy through different forms. Table 1 presents the four conditions to which participants were randomly assigned during their registration, the robot gave either empathetic feedback to the participant or not. An example of empathetic verbal feedback is:"

我可以想像到，最重要嘅係保持樂觀。” (“I can imagine. It’s important to stay positive”). An example of non-empathetic verbal feedback is:” 好的。” (“okay”). In the music condition, the session ended with the robot playing music or not. The complete experiment for one participant consisted of 3 sessions, with at least 2 days in between, over a period of 2 weeks.

Table 1

Experimental design

Empathy	Talk empathy	Music empathy
Condition 1	Yes	Yes
Condition 2	Yes	No
Condition 3	No	Yes
Condition 4	No	No

To reach participants with the 2021-2022 when the social isolation policy being in force and to reach the young adults suffering from loneliness and negative emotions (Sauter et al., 2020), participants interacted with a videotaped robot on the web. A small financial reward was provided, and all participants were uninformed about the actual background and conditions of the study. During the online registration through social media, participants digitally signed their informed consent. The study was approved by the Ethics Committee of the Hong Kong Polytechnic University of Hong Kong (Application Number: HSEARS20210623005).

A total of 144 Cantonese speakers (M_{age}= 23.7, SD_{age}= 3.33, 65.3% female, 76.39% Hongkongers) completed at least one session of our experiment. The sampling size exceeded the minimum number calculated using G-Power (Faul, Erdfelder, Buchner, Lang, 2009) with effect size considered to be medium (Cohen’s, 1988), alpha and power were .05 and .80 (Field, 2013) respectively. The effect size was chosen according to the average value in the media studies (Vermeulen & Beukenboom， 2016). Four participants with severe acquiescence response-bias in all interactions were excluded from the dataset. Acquiescence response-bias is the tendency to agree with each statement regardless of its content. After that, we had 140 participants engaged in the first session, 130 in the second session and 128 in the third one. Data was preserved from all participants that followed the proper sequence of sessions. For someone who followed session 1, skipped 2, and did do session 3, only the data of session 1 was kept. Therefore, participant #20 was removed from three sessions as s/he skipped the interaction video in the first session. Participants #50, #62, #68, #83 joined the third session but did not follow the second and their data were removed from the third session. Likewise, participant #36, #39 were removed from the third session as they skipped watching the third video (and so underwent no treatment). Consequently, there were 139, 129 and 121 participants left in the three waves, respectively.

Procedure

Upon clicking a link sent through email at each session, participants were instructed to fill out the first part of the questionnaire about their mental state, which was followed by the interaction with the robot in the video, after which they assessed their experience of the robot interaction. In the videos, the robot sat on the table and spoke to the participant, at times nodding its head, waving, and opening its arms. The interaction videos (between 40s to 90s) comprised of a greeting, a question-and-answer interaction, and an epilogue. To homogenize the interaction experience and for the simplicity of methods, though no time limit was imposed on the participants for the interactions and the questionnaire, the participants could not skip or review any robot questions, nor step forward and backward during the video playback. Participants were limited to select a single-choice response to the communications of the robot. The robot then offered empathetic or neutral feedback that further advanced the dialogue. The interaction protocols were designed based on the conversational models of Rashkin et al. (2018) and are presented in Appendix 9. Because the robot had no Cantonese language library, we dubbed the robot videos into Cantonese through the Google text-to-speech service and our own mastery of Cantonese.

Apparatus and Materials

Robot

The robot in the interaction video is NAO Humanoid Robot (see Figure 2) manufactured by Softback Robotics. It is a machine that is 58 cm tall and includes a variety of functions. It is also equipped with a wide range of sensors and controls, including a linear inverse pendulum, two cameras, and a sonar rangefinder, amongst others. In addition to that, it possesses a voice synthesizer, an inertial board, and a processor with an Intel ATOM 1.6 GHz clock speed. It is estimated that its battery life is greater than 1.5 hours (SoftBank Robotics, 2019). Choregraphe, the software development kit offered by Softbank Robotics, is utilized during the construction of the NAO. It has become one of the most common robotic platforms used in research and education and thus considered to be an appropriate model on which to focus our human–robot interaction study.

Figure 2: Nao Robot Created by Softbank Robotics

Online Interaction Platform

We made the interaction videos via Mindstamp[1], which is the easy-to-use interactive video creation platform. It allows us to ask questions directly during in the video that are free-

[1] https://mindstamp.io/

response, single-choice, multiple-choice, audio responses, etc. Importantly, we can get a full report showing exactly what the viewer did during the video playing, with which we can exclude those who skipped the video and the interaction according to the view and the response record. Each video is comprised of a greeting, a question-and-answer interaction, and an epilogue lead by the Nao in the video. The duration of the video varies from 40s to 90s. The video composed of the empathetic response and the music playing component should takes longer time. Please check the Appendix 10 for the interaction scripts.

Measurement

The questionnaire consisted of 56 Likert-type items, each followed by 6-point rating scales (1 = totally disagree; 6 = totally agree). Sixteen items measured the mental state of the participants before the interaction, 39 queried the feeling about the experience, and we inserted one manipulation check. The complete questionnaire covered 11 measurement scales, which are described per construct next.

Mental state consisted of two scales: Loneliness and Emotion, where Emotion was divided into Positive Emotion and Negative Emotion. For Loneliness, we selected a total of 6 items from De Jong Gierveld and Van Tilburg (2008), one example of which was “I experience a general sense of emptiness. (我經歷了巨大的空虛)” We added two items querying the respondents’ experience of loneliness during COVID-19. For Emotion, we used 4 positive emotion words (i.e., happy, relaxed, fulfilled, pleasant) and 4 negatives (i.e., sad, anxious, disgusted, fearful).

The **experiential variables** were measured by eight measurement scales, each indicating one of eight perceptions elaborated as follows: Bonding, Perceived relevance, Valence, Perceived empathy, Experienced empathy, Perceived realism, and Emotion afterwards.

Bonding was measured with four indicators (i.e., good feeling, fondness, connected, bonded) and two counter-indicators (i.e., as a stranger, cold). Two examples of indicators were “I did bond with Zora” and “I did have the feeling of contact with Zora.” These items were based on Broersen (2018) and Konijn and Hoorn (2005).

Perceived relevance was based on Van Vugt et al. (2006) and consisted of four items (need, important, help, meaningful). Two examples were “To me Zora has meaning” and “Zora ties in with what I need now.”

Valence was elaborated in terms of ‘below expectation’ and ‘beyond expectation,’ constructed from four items (positive expectation met, positive expectation unmet, negative expectation met, negative expectation unmet). Examples were “Zora’s performance was beyond my positive expectations” and “Zora’s performance was worse than I expected.”

Perceived Empathy measured whether the participants perceived empathy from the robot. It contained two indicators (understanding, appreciate) and two counter-indicators (incomprehension, mechanical). “I could see that Zora understood me” is an example of an indicator.

Experienced Empathy included two indicators (being understood, care) and two counter-indicators (being ignored, unaffected), measuring to what extent the participants were affected by the robot’s empathy. An example was “Zora’s response gave me a sense of being understood.”

Perceived realism was constructed from four items (real person, real conversation, real company, fake conversation). It measured the experienced realism rather than the evaluation of a realistic appearance (cf. Paauwe et al., 2015). Two examples were: “It feels like a real conversation to me” and “Zora feels like real company for me.”

Designed realism assessed the participants' evaluation of the human-likeness in the aspects of the robot's appearance and behaviours. It was constructed with four items based on TAB (Konijn et al., 2021), an example of which is "Zora looks like a real person."

Emotion afterwards was designed to know the psychological state (i.e., emotionally) of the participant, within a given situation (Haile, Gallagher & Robertson, 2015). We selected 7 most frequent emotional words we assumed people feel during the interaction. The scale consisted of 3 negative words, 3 positive words and a neutral word. According to Konijn and Hoorn (2003), affective realism causes us to experience supposed "realism" about the virtual figure that are in fact created by our feelings. In our study, *Emotion afterwards* scale was used to account for the value of *Perceived realism* on the robot.

To check the manipulation, participants also rated the overall level of empathy they experienced (single item) after treatment. Additional control variables were single items pertaining to age, gender, education, and country. Table 1 presents all the items on the questionnaire along with their abbreviations of scale names.

Table 2

Questionnaire items and scale names. R = reverse-coded counter-indication.

Scale	Abbr.	Content
Loneliness	*Lon#_1*	I experience a general sense of emptiness. 我經常感覺到空虛。
	Lon#_2	I miss having people around. 我懷念身邊有人圍繞的感覺。
	Lon#_3	I often feel rejected. 我經常感受到別人的拒絕。
	Lon#_4-> Lon#_4R	There are plenty of people I can rely on when I have problems. 當我遇到問題的時候，我感覺我身邊有很多人可以讓我依靠。(R)
	Lon#_5-> Lon#_5R	There are many people I can trust completely. 我有很多完全值得我信賴的人。(R)
	Lon#_6-> Lon#_6R	There are enough people I feel close to. 我身邊有足夠多親近的人。(R)
	Lon#_7	I miss the fun around me. 我想念充滿歡樂的時光。
	Lon#_8	I miss the social connection. 我想念社交接觸。
Positive emotion	*Emo#_1*	Happy 開心
	Emo#_2	Relaxed 輕鬆
	Emo#_3	Satisfied 滿足
	Emo#_4	Joyful 喜悅
Negative emotion	*Emo#_5*	Sad 傷心 (R)
	Emo#_6	Anxious 焦慮 (R)
	Emo#_7	Disgust 厭惡 (R)

	Emo#_8	Fear 害怕 (R)
Bonding	*Bon#_1*	Zora made me feel good. Zora 使我感覺很好。
	Bon#_2	I liked it with Zora. 我喜歡跟 Zora 在一起。
	Bon#_3	I did bond with Zora. 我與 Zora 有建立關係。
	Bon#_4	I did have the feeling of contact with Zora. 我感覺到我和 Zora 有交流。
	Bon#_5	I liked Zora. 我喜歡 Zora。
	Bon#_6	I would like to repeat this experience. 我想要再次經歷這段體驗。
	Bon#_7 -> Bon#_7R	For me Zora remains a stranger. 對我來說，Zora 仍然是一個陌生人。(R)
	Bon#_8 -> Bon#_8R	Zora remains a cold device for me. 對我來說，Zora 仍然是一個冷漠的機器。(R)
Perceived empathy	*PerEmp#_1*	I could see Zora showing understanding for me. 我能夠看到 Zora 在對我表示理解。
	PerEmp#_2	Zora appreciated exactly how the things I experienced felt to me. Zora 非常認同我的感覺。
	PerEmp#_3	Zora’s response to me is without understanding. Zora 沒有帶著理解來回應我。(R)
	PerEmp#_4	No matter what I tell about myself, Zora acts just the same. 無論我說什麼，Zora 的反應 都是相同的。(R)
Experienced empathy	*EpthyHum#_5*	Zora’s response gave me a sense of being understood. Zora 讓我有一種被理解的感覺。
	EpthyHum#_6-> EpthyHum#_6R	It was difficult for me to share common feelings with Zora(R). 我沒辦法對 Zora 感同身受。
	EpthyHum#_7	I cared about Zora’s affective expression in its response. 我關心 Zora 在回應中所表達出來的情感。
	EpthyHum#_8-> EpthyHum#_8R	I remained unaffected no matter Zora’s response to me(R). Zora 的回應沒有讓我有任何情緒波動。
Perceived relevance	*Rel#_1*	I do need Zora. 我需要 Zora。
	Rel#_2	It's important to me to have Zora. 擁有 Zora 對我說是重要的。
	Rel#_3	To me Zora has a lot of meaning. 對我來說，Zora 具有很多意義。
	Rel#_4	Zora ties in with what I need now. Zora 對滿足我現在的需求是有關係的。
Valence	*E#pect#_1*	Zora met my initial expectations. Zora 達到了我的預期。
	E#pect#_2-> E#pect#_2R	I had a higher expectation on Zora. 我對 Zora 的期望過高了。(R)
	E#pect#_3	I had a lower expectation on Zora. 我對 Zora 的期望過低了。
	E#pect#_4-> E#pect#_4R	Zora failed to meet my initial expectation. Zora 並未達到我的期望。(R)
Experienced realism	*E#pReal#_1*	It’s like talking to a real person. 我好像在跟一個真人對話。
	E#pReal#_2	It feels like a real conversation to me. 剛才的對話很真實。
	E#pReal#_3	Zora feels like real company for me. Zora 給了我真實的陪伴。
	E#pReal#_4-> E#pReal#_4R	It differed from a real conversation. 剛才的對話有別於真實的對話。(R)
Design realism	*DesReal#_1*	Zora looks like a real person. Zora 看起來像真人。
	DesReal#_2	Zora’s appearance looks human. Zora 的外觀看起來像人。
	DesReal#_3	Zora acts like a human. Zora 的行為像人。
	DesReal#_4-> DesReal#_4R	Zora is more of a machine. Zora 更像機器。
Emotion afterwards	*EmoAfter#_1*	Interested 感興趣
	EmoAfter#_2	Happy 開心
	EmoAfter#_3	Surprised 驚訝
	EmoAfter#_4	Calm 平靜
	EmoAfter#_5 -> EmoAfter#_5R	Boring 無聊 (R)
	EmoAfter#_6 -> EmoAfter#_6R	Disappointed 失望 (R)
	EmoAfter#_7 -> EmoAfter#_7R	Fear 害怕 (R)
Manipulation check	*EpthyCheck*	To what extent do you agree that Zora’s response was empathic?

		你有多認同這個表述: Zora 的回應是富有同理心的?
Demographic details	*Gender*	性別
	Age	年齡
	Country	城市
	Education	教育程度

*# represents the session number.

Before doing reliability analysis, we reverse-coded (1→6, …, 6→1) the counter-indicative items across three sessions for three Loneliness items (*Lon#*_5R, *Lon#*_6R, *Lon#*_7R), two Bonding items (*Bon#*_7R, *Bon#*_8R), two Experienced empathy items (*EpthyHum#*_6R, *EpthyHum#*_8R), two Valence items (*Expect#*_2R, *Expect#*_4R), one Experienced realism (*ExpReal#*_4R) and Designed realism respectively (*DesReal#*_4R), and three Emotion afterwards (*EmoAfter#*_5R, *EmoAfter#*_6R, *EmoAfter#*_7R).

Subsequently, to test whether the measurement scale achieved good reliability, we calculated Cronbach's Alpha on scales with the items within its scale, after which we ran Principal Component Analysis (PCA) to extract factors with significant eigenvalues.

Reliability Analysis

Reliability is defined as consistency in results after repeated measurements (Taylor, 1990). This consistency is defined by the homogeneity of the results. In our study, we would like to check the homogeneity among the items within the same scale. Thus, we examined this understanding consistency across participants. We argued that if the participants understood the questionnaire as we expected in their first exposure to it, they should show at least an equal understanding in the second and third exposure. Thus, we established measurement reliability on the observations in the first session.

Cronbach's α coefficient values were assessed (Cronbach, 1951) in order to measure the internal reliability of each construct. For the mental state scales, Lon1 scale could be improved to the conventional cut-off point of 0.7 (Nunnally, 1975), scoring Cronbach's $\alpha = .74$, after removing *Lon1_2*, *Lon1_7* and *Lon1_8*. *Emo1_n* achieved highest reliability (Cronbach's α = .74) with 4 items in. Although the items of *Emo1_p* achieved good reliability (Cronbach's α > .79), it could be improved by removing *Emo1_2*. The content of *Emo1_2* was "relaxed" which is more neutral compared to the other three adjectives (e.g, happy, joyful). To strengthen the characteristics of this scale, we decided to exclude *Emo1_2* and enhanced the reliability to Cronbach's α = .88. The scales *Bon1*, *PerEmp1*, *EpthyHum1*, *Rel1*, *DesReal1*, with all items included achieved good to very good reliability in the first run (.92 < Cronbach's α > .79). *ExpReal1* increased to Cronbach's α = .86 after removing *ExpReal1_4R*. Expect1 scored Cronbach's α = .75 with all items in and could not be improved by eliminating any item. Results are compiled in Table 3.

Table 3

Reliability of scales (data from first session)

Scale	Num of items	Items	Alpha / r	Standardized Alpha	Item mean	Variances	*N*
Lon1	5	*Lon1_1, Lon1_3, Lon1_4R, Lon1_5R, Lon1_6R* (*Lon1_2, Lon1_7, Lon1_8* removed)	.75	.75	3.02	1.25	139
Emo1_p	3	*Emo1_1, Emo1_3, Emo1_4* (*Emo1_2* removed)	.88	.89	3.75	1.16	139
Emo1_n	4	*Emo1_5 ~ Emo1_8*	.81	.82	4.09	1.60	139
Bon1	8	*Bon1_1 ~ Bon1_6, Bon1_7R, Bon1_8R*	.90	.90	3.38	1.11	139
PerEmp1	4	*PerEmp1_1, PerEmp1_2, PerEmp1_3R, PerEmp1_4R*	.79	.79	3.40	1.32	139
EpthyHum1	4	*EpthyHum1_5, EpthyHum_6R, EpthyHum_7, EpthyHum_8R*	.79	.80	3.19	1.26	139
Rel1	4	*Rel1_1~Rel1_4*	.92	.93	2.69	1.25	139
Expect1	4	*Expect1_1, Expect_2R, Expect_3, Expect_4R*	.75	.75	3.46	1.23	139
ExpReal1	3	*ExpReal1_1, ExpReal1_2, ExpReal1_3* (*ExpReal1_4R* removed)	.86	.86	2.85	1.32	139
DesReal1	4	*DesReal1_1, DesReal1_2,*	.83	.83	2.38	1.08	139

		DesReal1_3, *DesReal1_4R*					
EmoAfter1	3	*EmoAfter1_1*, *EmoAfter1_2*, *EmoAfter1_3*, *EmoAfter1_5R* (*EmoAfter1_4*, *EmoAfter1_6R*, *EmoAfter1_7R*)	.74	.74	3.46	1.39	149

Validity analysis

After reliability analysis, two principal component analyses (PCAs) were conducted to examine the mental state factor structure and the experience factor structure of the questionnaire respectively with their remaining items. The PCA used Promax rotation with free fitting format and Kaiser normalization. Extraction and retention of factors was based on visual examination of the scree plot (Cattell, 1966) and eigenvalues of >1.0 were retained (Kaiser, 1960). The threshold for the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy was 0.5 (Tabachnick & Fidell, 2001). A factor loading threshold of .50 was applied to enhance the strength of factors, so only items of this strength, or with loadings that could be rounded up to .50, were retained. For factor loadings of .5 to be significant, a minimum of 120 participants are required to allow sufficient statistical power (Hair et al., 2019, page 116). The sample sizes of three session were sufficient to meet this threshold.

Table 4 showed the PCA result with experiential items in first session. 7 component loadings in a fitting format were found. The Kaiser-Meyer-Olkin measure of Sampling Adequacy (KMO =0.92) and Bartlett's test of sphericity [$X^2(595) = 3562.14$, $P \leq 0.001$] indicated that the sample size was adequate (Field, 2014; Kaiser, 1974) and the data were appropriate (there is a significant difference in the variances) for conducting PCA. The PCA extracted seven significant factors that explained 70.71% of the total variance. The results turned out that four items (*Bon1_1*, *Bon1_2*, *Bon1_5*, *Bon1_6*) within *Bon1* scale clustered into the first component, whereas all *Rel1* items (*Rel1_1* ~ *Rel1_4*) and most of *PerEmp1* items (*PerEmp1_1*, *PerEmp1_2*, *PerEmp1_3R*) were placed in component 2 and component 3, respectively. Concerning *DesReal1*, we found its items (*DesReal1_1*, *DesReal1_2*, *DesReal1_4R*) were convergent into component 4 and two items of *Expect1* (*Expect1_2R*, *Expect1_4R*) into component 5. The items of *EpthyHum1*, *ExpReal1* and *EmoAfter1* were spread over the components or with lower loading to any component. Yet, *EpthyHum1* and *ExpReal1* were theoretically interesting in our study. By contrast, *EmoAfter1* was less important as it was used to account for the variable of *ExpReal1*. Given the items of this scale were placed at the end of the questionnaire and would not influence the previous observations, we decided to rerun PCA after removing EmoAfter1 items. Component loadings and item statistics are presented in Table 5. The KMO (≥.93) indicated that now the data were more appropriate for conducting PCA (Field, 2014; Kaiser, 1974). Also, the component extractions were neat. *PerEmp1_4R* was clustered into *PerEmp1* scale, and *DesReal1_3* into *DesReal1* scale. Thus, we would follow the second PCA results.

Table 4

Pattern Matrix with 7 components on experiential variables (first run, first session)

	Component						
	1	2	3	4	5	6	7
Bon1_1	.598		.326				
Bon1_2	.755						
Bon1_3	.391		.349	.444			
Bon1_4	.322	-.308	.592	.312			
Bon1_5	.673		.362				
Bon1_6	.716						
Bon1_7R					.784		-.409
Bon1_8R				.467	.430		
PerEmp1_1			.959				
PerEmp1_2			.948				
PerEmp1_3R			.511			.345	
PerEmp1_4R	-.322		.438		.644		
EpthyHum1_5			.548				
EpthyHum1_7		.300	.317				
EpthyHum1_8R					.961		
EpthyHum1_6R			.319		.398		
Rel1_1		.931					
Rel1_2		.967					

Rel1_3		.707					
Rel1_4		.951					
Expect1_1			.571				
Expect1_3	.583					.325	
Expect1_2R						1.009	
Expect1_4R						.779	
ExpReal1_1							
ExpReal1_2						.374	
ExpReal1_3		.359					
DesReal1_1		.326		.788			
DesReal1_2		.317		.767			
DesReal1_3				.482			
DesReal1_4R				.786			
EmoAfter1_1	.524						.407
EmoAfter1_2	.331				.388		.343
EmoAfter1_3							.970
EmoAfter1_5R	.559					.420	

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalization.

a. Rotation converged in 14 iterations.

Table 5

Pattern Matrix with 6 components on experiential variables (second run, first session)

	Component					
	1	2	3	4	5	6
Bon1_1		.669				
Bon1_2		.785				
Bon1_3		.469		.458		
Bon1_4	-.308	.407	.534	.309		
Bon1_5		.684	.337			
Bon1_6		.707				
Bon1_7R					.755	
Bon1_8R				.445	.395	
PerEmp1_1			1.004			
PerEmp1_2			.922			
PerEmp1_3R			.595			
PerEmp1_4R			.519		.547	
EpthyHum1_5			.468			
EpthyHum1_7	.404		.352	-.311		
EpthyHum1_8R					.986	
EpthyHum1_6R					.358	
Rel1_1	.895					
Rel1_2	.919					

Rel1_3	.733					
Rel1_4	.903					
Expect1_1			.536			
Expect1_3		.577				.365
Expect1_2R						1.020
Expect1_4R						.815
ExpReal1_1				.308		
ExpReal1_2			.351	.354		.325
ExpReal1_3	.418					
DesReal1_1				.830		
DesReal1_2				.793		
DesReal1_3				.539		
DesReal1_4R				.838		

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalization.

a. Rotation converged in 7 iterations.

Then we examined the factor structure of the mental state items. The KMO test statistic was .86, considered as meritorious and the sampling was adequate. Bartlett's test of sphericity was also significant [$X^2(66) = 724.75$, $P \leq 0.001$], indicating that the individual correlations within the R-matrix of the PCA were sufficient for interpretation. The pattern matrix was chosen to report the factor loading about questions (Table 6). We identified three distinct components, which totally accounted for 66.13% of the total variance. The second component consisted of 3 items (*Emo1_1*, *Emo1_3*, *Emo1_4*) of *Emo1_p*. The majority of *Lon1* items (*Lon1_4R*, *Lon1_5R*, *Lon1_6R*) fell on the third component. Except for the four *Emo1_n* items (*Emo1_5*, *Emo1_6*, *Emo1_7*, *Emo1_8*), *Lon1_1* with significant loading value (.653) contributed to component 1 as well. Though the statement of *Lon1_1* suggested emptiness, this negative word represented a complex emotion and may have messed up the characteristic of the basic negative emotion scale. We decided to remove this item from the first component. Next, we re-ran the reliability analyses to see whether the remaining items well composed its scale.

Table 6

Pattern Matrix with 3 components on mental states variable (first run, first session)

	Component		
	1	2	3
Lon1_1	.653		
Lon1_3	.443		.468
Lon1_4R			.827
Lon1_5R			.844
Lon1_6R			.718
Emo1_1		.918	
Emo1_3		.815	
Emo1_4		.867	
Emo1_5	.756		
Emo1_6	.687		
Emo1_7	.757		
Emo1_8	.898		

Extraction Method: Principal Component Analysis.
Rotation Method: Promax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Reliability Analysis after PCA

In all, we had three mental-state scales and four experiential scales that achieved good to very good reliability in their separate run (*Lon1*: Cronbach's α = .79; *Emo1_p*: Cronbach's α = .88; *Emo1_n*: Cronbach's α = .81; *Bon1*: Cronbach's α = .88; *PerEmp1*: Cronbach's α = .79; *Rel1*: Cronbach's α = .93; *DesReal1*: Cronbach's α = .83). There were only two items in *Expect1* scale, thus we calculated Spearman-Brown Correlation. *Expect_2R* and *Expect_4R* were significantly correlated (r = .62***). Table 7 offers a summary.

To check whether the results were reproducible across time, we also performed the reliability analysis with the data from the second and third session. Compared with Table 6, 7, and 8, we found the reliability was not only replicated but even improved.

Table 7

Reliability of scales (first session, after PCA)

Scale	Num of items	Items	Alpha / r	Standardized Alpha	Item mean	Variances	*N*
Lon1	3	*Lon1_4R*, *Lon1_5R*, *Lon1_6R*	.79	.79	2.88	1.28	140
Emo1_p	3	*Emo1_1*, *Emo1_3*, *Emo1_4* (*Emo1_2* removed)	.88	.89	3.74	1.17	140
Emo1_n	4	*Emo1_5 ~ Emo1_8*	.81	.81	2.91	1.59	140
Bon1	4	*Bon1_1*, *Bon1_2*, *Bon1_5*, *Bon1_6*	.88	.89	3.64	1.09	140
PerEmp1	4	*PerEmp1_1*, *PerEmp1_2*, *PerEmp1_3R*, *PerEmp1_4R*	.79	.79	3.40	1.31	140
Rel1	4	*Rel1_1~Rel1_4*	.93	.93	2.68	1.26	140
Expect1	2	*Expect_2R*, *Expect_4R*	r = .62**				140
DesReal1	4	*DesReal1_1*, *DesReal1_2*, *DesReal1_4R*, *DesReal1_3*	.83	.83	2.37	1.08	140

Table 8

Reliability of scales (second session, after PCA)

Scale	Num of items	Items	Alpha / r	Standardized Alpha	Item mean	Variances	*N*
Lon2	3	*Lon2_4R*, *Lon2_5R*, *Lon2_6R*	.90	.90	2.97	1.04	129
Emo 2_p	3	*Emo2_1*, *Emo2_3*, *Emo2_4*	.93	.94	3.75	1.14	129
Emo 2_n	4	*Emo2_5 ~ Emo2_8*	.85	.86	2.73	1.43	129
Bon 2	4	*Bon2_1*, *Bon2_2*, *Bon2_5*, *Bon2_6*	.93	.94	3.40	1.09	129
PerEmp 2	4	*PerEmp2_1*, *PerEmp2_2*, *PerEmp2_3R*, *PerEmp2_4R*	.81	.81	3.24	1.23	129
Rel 2	4	*Rel2_1~Rel2_4*	.95	.96	2.66	1.17	129

Expect 2	2	*Expect_2R, Expect_4R*	*r* = .61**				129
DesReal2	4	*DesReal2_1, DesReal2_2, DesReal2_4R, DesReal2_3*	.89	.90	2.44	1.13	129

Table 9

Reliability of scales (third session, after PCA)

Scale	Num of items	Items	Alpha / r	Standardized Alpha	Item mean	Variances	*N*
Lon3	3	*Lon3_4R, Lon3_5R, Lon3_6R*	.92	.92	2.99	1.11	121
Emo 3_p	3	*Emo3_1, Emo3_3, Emo3_4*	.90	.90	3.75	1.27	121
Emo 3_n	4	*Emo3_5 ~ Emo3_8*	.84	.85	2.65	1.47	121
Bon_3	4	*Bon3_1, Bon3_2, Bon3_5, Bon3_6*	.94	.94	3.42	1.26	121
PerEmp_3	4	*PerEmp3_1, PerEmp3_2, PerEmp3_3R, PerEmp3_4R*	.83	.83	3.29	1.27	121
Rel_3	4	*Rel3_1~Rel3_4*	.96	.96	2.79	1.51	121
Expect_3	2	*Expect_2R, Expect_4R*	*r* = .71**				121
DesReal_3	4	*DesReal3_1, DesReal3_2, DesReal3_4R, DesReal3_3*	.89	.89	2.52	1.26	121

Exploration of outliers

We then calculated the means across the items on the 8 remaining scales × 3 sessions. The 24 mean values are distinguished from the single-item values by the prefix *M_*. Next, we did outlier analyses on each scale via Boxplots with the factors of *MusicEmpathy* and *TalkEmpathy*. Boxplots indicated some potential outliers for all 24 variables. For *M_PerEmp1*, participant #44 scored values outside the 3rd interquartile range and was eliminated from the first session. We reran the outlier analysis on the variables in the first session and ended up with the outlier distribution in Table 10.

In Table 10, outliers in the lower end of the box are marked with a bracket. In itself, outliers, are legitimate observations that are a natural part of the population, and we had no hypotheses to exclude them. Yet, outliers can distort statistical analyses and violate model assumptions, so we performed our analyses with and without outliers included.

We extracted the outliers per session (see Table 11, 12, and 13). As a result, we had two data sets for session 1: with outliers (*N1* = 139) and without (*n1* = 125). Likewise, there were two data sets for session 2 (*N2* = 129, *n2* = 119) and session 3 (*N3* = 121, *n3* = 112).

In addition, we would like to see whether the mental-state outliers disrupted the average experience. Most of them were in a negative state, except for participant 80 in session 2. Therefore, we categorized the outliers into mental-state outliers and experiential outliers to inspect them separately.

Table 10

Outlier distribution over three sessions

Condition	**1**			**2**			**3**			**4**		
Session	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
M_Lon	6			58,59		58,59				138,139		
M_EmoP	(15,17,31)	(15,18)			(37)	(59)	(94)	(78),80				
M_EmoN	17								93		108,120	108
M_Bon		(18)	(16)			(57,59,65)					(108)	
M_PerEmp										(134)		
M_Rel												
M_Expect	21		9	34	(59)	59						
M_DesReal	27	27		63								116

* Outliers in the lower end of the box are marked with a bracket

Table 11

Between-Subject outlier distribution (first session)

Items	**Con 1**	**Con2**	**Con3**	**Con4**
M_Lon1	6	58,59	/	138,139

M_EmoP1	(15,17,31)	/	(94)	/
M_EmoN1	17	/	/	/
M_Bon1	/	/	/	/
M_PerEmp1	/	/	/	(134)
M_Rel1	/	/	/	/
M_Expect1	21	34	/	/
M_DesReal1	27	63	/	/
Mental-state outliers	6,15,17,31	58,59	94	138,139
Experiential outliers	21,27	34,63		134
All outliers	6,15,17,21,27,31	34,58,59,63	94	134,138,139

* Outliers in the lower end of the box are marked with a bracket

Table 12

Between-Subject outlier distribution (second session)

Items	Condition 1	Con2	Con3	Con4
M_Lon2	/	/	/	/
M_EmoP2	(15,18)	(37)	(78),80	/
M_EmoN2	/	/	/	108,120
M_Bon2	(18)	/	/	(108)
M_PerEmp2	/	/	/	/
M_Rel2	/	/	/	/
M_Expect2		(59)		
M_DesReal2	27			116
Mental state outliers	15,18	37	78,80	108,120
Experiential outliers	18,27	59		108,116
All outliers	15,18,27	37,59	78,80	108,116,120

* Outliers in the lower end of the box are marked with a bracket

Table 13

Between-Subject outlier distribution (third session)

Items	Condition 1	Con2	Con3	Con4
M_Lon3	/	58,59	/	/
M_EmoP3	/	(59)	/	/
M_EmoN3	/	/	95	108
M_Bon3	(16)	(57,59,65)	/	/
M_PerEmp3	/	/	/	/
M_Rel3	/	/	/	/
M_Expect3	9	59		
M_DesReal3				116
Mental state outliers		58,59	95	108
Experiential outliers	9,16	57,59,65		116
All outliers	9,16	57,58,59,65	95	108,116

* Outliers in the lower end of the box are marked with a bracket

Results

Demographics

Chi-square tests were conducted to check whether random assignment of participants was successful. One assumption of Chi-square is all the variables should be categories. We recoded the continuous *Age* into categorical *Age range* (checked Table 14). Table 15 reports the results of Chi-square tests of the null hypothesis that participants were randomly distributed across conditions along four different observable participant characteristics: *Gender*, *Education*, *Age range*, and *Country*. The tests were based on data from session 1 with outliers (*N1* = 139); significance was based on the 5-percentage level. Before we looked at the significant values of the tests, we checked whether the data met the Pearson Chi-Square assumptions (Table 15, Column 5). For each demographics variable, if the result percentage is greater than 20%, then the assumption of the data within this variable has been violated and we needed to take a different course of action. From the last column of Table 15, we found only the test with *Country* did not violate the assumptions. When we read its Pearson Chi-square value, however, it said 2.01 with .57 (χ^2 = 2.01, p = .57) asymptotic significance (≥.10). This statistical non-significance enabled us to accept our null hypothesis, which stated that there was no significant association between country and condition status. In other words, country was independent from condition.

Regarding *Gender*, *Education* and *Age range*, we further read off their likelihood Ratios (checked lower part of Table 15). It showed not statistically difference of the observations between conditions in terms of the characteristics of gender (χ^2 = 2.47, p = .87), education (χ^2 = 6.67, p = .35) and age range (χ^2 = 8.48, p = .93). In conclusion, the demographic comparison across conditions showed observations were randomly and independently sampled from the population, and thus we could conduct the Multivariate analyses of variance (MANOVA) without adding them as covariate.

Table 14

Demographic distribution over experimental conditions (N1 = 139)

	Condition 1 (*n*= 32)	Condition 2 (*n*= 37)	Condition 3 (*n*= 33)	Condition 4 (*n*= 37)	Overall (*N*= 139)

Gender					
Prefer not to say	2	2	4	3	11
Female	20	23	21	26	90
Male	10	12	8	8	38
Education					
Less than high school	0	0	0	0	0
High school	1	2	0	0	3
College	6	3	4	7	20
University	25	32	29	30	116
Age (years)					
20 and below	3	5	4	6	18
21-25	22	22	18	21	83
26-30	6	10	8	10	34
Above 30	1	0	3	0	4
Country					
Mainland	5	10	7	10	32
Hong Kong	27	27	26	27	107

Table 15

Pearson Chi-Square Value on Age, Gender, Country, and Education across Conditions (N1 = 139)

	Value	Df	Asymptotic Significance (2-sided)	Pearson Chi-Square Assumption *
Gender * Condition	2.53	6	.87	Violated (33.3%)
Country * Condition	2.01	3	.57	Not violated (.0%)
Education * Condition	5.66	6	.46	Violated (50%)
Age Range * Condition	7.98	9	.54	Violated (50%)
			(Pearson Chi-Square)	
Gender * Condition	2.47	6	.87	
Education * Condition	6.67	6	.35	
Age Range * Condition	8.48	9	.93	
			(Likelihood Ratio)	

* The assumption is what percentage of the cells have expected count less than five.

We ran the Chi-Square tests with the other session data and the results suggested that the participants were successfully assigned to groups randomly.

Manipulation Checks

We ran 24 one-sample *t*-tests (two-tailed) over 4 conditions × 3 sessions × 2 data sets (with and without outliers) with 3 (‘disagree a little’) as the test value to see whether the participants recognized the empathic behaviours performed by Zora as we expected. Table 16 shows the results. The *t*-values in Condition 1 were positive, which gave us 2-tailed significance values below .05. This may count as evidence that the participants in Condition 1 experienced empathy from the talking and music-playing Zora, which met our aim.

Results in Condition 2 showed significant difference between the sample group and the population group, expect for those (with outliers) who participated in session 3 (t =1.79, p < .083). However, the difference became significant after removing the outliers (t =2.15, p < .04). Regarding Condition 3 and 4, the participants only reported the significant perceived empathy from Zora (with test value =3) in their first exposure to the empathic robot. The perception of empathy from Zora went away in the second and third exposures. Remarkably, mean values in session 2 (M = 2.86, SD =1.30 with outliers) and 3 (M = 2.97, SD =1.27 without outliers) suggested that less empathy was perceived by the participants when Zora only played music.

When we inspected the differences across sessions, the results suggested that in the first session, all participants experienced empathy from Zora through music and speech. However, we had no evidence whether that perception came from the topic in first session or novelty effects. Nevertheless, participants seemed to perceive empathy through music and talking, and regarded Zora least empathetic if it only played music.

Table 16

One-sample t-tests (3 is the test value), checking whether robot empathy was perceived after interaction

Condition		Mean	Std. Deviation	t	Sig. (2-tailed)	N
MusicEmpathy - TalkEmpathy	EpthyCheck1_1 (with outliers)	3.91	1.06	4.84	.000	32
	EpthyCheck1_1 (without outliers)	3.73	1.04	3.58	.001	26
	EpthyCheck2_1 (with outliers)	3.45	.96	2.62	.014	31
	EpthyCheck2_1 (without outliers)	3.50	.84	3.15	.004	28
	EpthyCheck3_1 (with outliers)	3.74	1.21	3.41	.002	31
	EpthyCheck3_1 (without outliers)	3.75	1.11	3.58	.001	28
MusicNo - TalkEmpathy	EpthyCheck1_1 (with outliers)	3.64	1.10	3.49	.001	36
	EpthyCheck1_1 (without outliers)	3.48	1.06	2.54	.016	31
	EpthyCheck2_1 (with outliers)	3.64	1.06	3.20	.003	28
	EpthyCheck2_1 (without outliers)	3.73	.96	3.88	.001	26
	EpthyCheck3_1 (with outliers)	3.36	1.17	1.79	.083	33
	EpthyCheck3_1 (without outliers)	3.45	1.12	2.15	.040	29
MusicEmpathy - TalkNo	EpthyCheck1_1 (with outliers)	3.38	.94	2.25	.032	32
	EpthyCheck1_1 (without outliers)	3.35	.95	2.08	.046	31

	EpthyCheck2_1 (with outliers)	2.86	1.30	-.57	.57	29
	EpthyCheck2_1 (without outliers)	2.89	1.32	-.43	.67	28
	EpthyCheck3_1 (with outliers)	2.97	1.27	-.15	.89	29
	EpthyCheck3_1 (without outliers)	3.04	1.23	.15	.88	28
MusicNo – TalkNo	EpthyCheck1_1 (with outliers)	3.67	1.01	3.94	.000	36
	EpthyCheck1_1 (without outliers)	3.67	1.05	3.64	.001	33
	EpthyCheck2_1 (with outliers)	3.20	.87	1.36	.18	35
	EpthyCheck2_1 (without outliers)	3.33	3.22	.87	.17	32
	EpthyCheck3_1 (with outliers)	3.31	1.03	1.72	.096	32
	EpthyCheck3_1 (without outliers)	3.30	.99	1.66	.11	30

Effects of Media on experiential variables

The robot showed empathy, using two Media: Music and Talk (**IVs**). To study the effect of Media on participants' experience, we conducted Multivariate Analyses of Variance (MANOVA), checking the data for model assumptions: 1) without multivariate outliers, 2) for homoscedasticity, 3) for absence of multicollinearity, and 4) independence and randomness (Salkind, 2010). For MANCOVAs, we also checked 5) the significant linear relationship between the dependent variables (**DVs**) and the covariate (Siri et al., 2018). Yet, some researchers believe it is not essential that a covariate is related to each DV in a MANOVA. If the covariate is related to none of the DVs, and the groups do not differ in average covariate value, then the inclusion of such a covariate would not affect the results, but it does no good either. In the **Demographics** section, we found no differences for demographic variables across conditions. To simplify analysis, we excluded all the demographic variables (*Gender*, *Age range*, *Country*) as covariates in MANOVA.

Some studies indicated that lonely people are likely to engage in robot interaction (Cyranowski et al., 2013; Waytz & Epley, 2012). To reduce error terms and eliminate the covariates' effect on the relationship between Media and DVs, we used the mental-state variables (*M_Lon*, *M_EmoP*, *M_EmoN*) as control variables in the MANOVA analyses.

Session 1

MANOVA on experiential variables with *N1* = 139. With *N1* = 139, we ran Pearson correlation analyses with independent variables (*M_Bon1*, *M_PerEmp1*, *M_Rel1*, *M_Expect1*, *M_DesReal1*). Table 17 shows the correlation coefficients. Between every two dependent variables, the coefficients were between .08 and .78, and thus there was no multicollinearity between the experiential variables ($|r| < 0.9$) (Tabachnick and Fidell, 2012). With regard to the multivariate outliers, we calculated the Mahalanobis distance on experiential variables with *N1* = 139 and related the results to the table of Critical Values (Schoen et al., 2011). As there were five dependent variables in our sample, the corresponding critical value was 20.52. The maximum Mahalanobis distance in our sample *N1* was 18.51 < 20.52, so there was no multivariate outlier in *N1* sampling.

With the assumption 1 and 3 met, we conducted MANOVA for the experiential variables with *N1* = 139, using mental-state variables as covariates. In the Levene's Test of Equality of Error Variances table of the output, the assumption of equal variances was confirmed as the *p*-values were all greater than 0.05 (*M_Bon1*: $F_{(3,\ 135)} = 1.84$. $p = .14$; *M_PerEmp1*: $F_{(3,\ 135)} = .96$. $p = .42$; *M_Rel1*: $F_{(3,\ 135)} = .53$. $p = .66$; *M_Expect1*: $F_{(3,\ 135)} = .50$. $p = .69$; *M_DesReal1*: $F_{(3,\ 135)} = 1.10$. $p = .35$). We checked the result of Box's Test of Equality of Covariance Matrices and found that $p > 0.001$, which indicated that the covariance in different conditions was significant. Therefore, we assumed that the dataset of *N1* =139 met the fourth model assumption of MANOVA. With all assumptions examined, we could study our results of testing the hypotheses.

Table 17

Correlation matrix for experiential variables and covariates (N1 =139)

Scale	M_Bon1	M_PerEmp1	M_Rel1	M_Expect1	M_DesReal1
		Condition 1			
M_Bon1	1				
M_PerEmp1	.562**	1			
M_Rel1	.700**	.557**	1		

M_Expect1	.720**	.546**	.617**	1	
M_DesReal1	.599**	.516**	.780**	.582**	1
Condition 2					
M_Bon1	1				
M_PerEmp1	.390*	1			
M_Rel1	.562**	.529**	1		
M_Expect1	.350*	.344*	.397*	1	
M_DesReal1	.600**	.524**	.695**	.195	1
Condition 3					
M_Bon1	1				
M_PerEmp1	.606**	1			
M_Rel1	.778**	.599**	1		
M_Expect1	.211	.446**	.302	1	
M_DesReal1	.640**	.573**	.750**	.186	1
Condition 4					
M_Bon1	1				
M_PerEmp1	.488**	1			
M_Rel1	.442**	.301	1		
M_Expect1	.399*	.595**	.084	1	
M_DesReal1	.458**	.480**	.518**	.428**	1

Note. Pearson correlation coefficients based on 5000 bootstrap samples are presented above (N =139).
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.000 level (2-tailed).

Wilks' Lambda statistics was used to examine whether there was an effect of Media on *MusicEmpathy* and *TalkEmpathy*. Table 18 shows that, after controlling for the mental-state variables, the effect of Media was not statistically significant (F=2.31, p=0.078, $Wilks'\ \Lambda$=0.93, $partial\ \eta^2$=0.083), indicating that the effect of music on participants' experience did not vary by talking or not. Also, the main effect of *MusicEmpathy* on the experiential variables was not statistically significant: F=0.71, p=0.62, $Wilks'\ \Lambda$=0.97; $partial\ \eta^2$=0.027. Neither was the main effect of *TalkEmpathy*: F=0.18, p=0.99, $Wilks'\ \Lambda$=.99; $partial\ \eta^2$=0.007. Covariates were significantly associated with a linear combination of scores on DVs (*M_Lon1*: F=2.74, p=0.022, $Wilks'\ \Lambda$=0.90, $partial\ \eta^2$=0.10; *M_EmoP1*: F=3.48, p=0.006, $Wilks'\ \Lambda$=0.90, $partial\ \eta^2$=0.12).

Table 18

Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (N1 =139)

Effect	**Value**	***F***	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.674	12.381	5.000	128.000	.000	.326
M_Lon1	.903	2.743	5.000	128.000	.022	.097
M_EmoP1	.880	3.480	5.000	128.000	.006	.120
M_EmoN1	.922	2.174	5.000	128.000	.061	.078
MusicEmpathy	.973	.709	5.000	128.000	.618	.027
TalkEmpathy	.993	.180	5.000	128.000	.970	.007
*MusicEmpathy * TalkEmpathy*	.926	2.307	5.000	128.000	.078	.083

Although we could not find the interaction effects of *MusicEmpathy* and *TalkEmpathy* on a combination of scores on DVs, we further explored if *Media* influenced any sole experimental variable. The Tests of Between-Subjects Effects in the result output (Table 19) shows the interaction of Media led to the significant difference on *M_Bon1*(F=7.22, p=0.008, *partial* η^2=0.52) and *M_PerEmp1*(F=6.91, p=0.010, *partial* η^2=0.050). Therein, *M_EmoP1* moderated the Media interaction effect on *M_Bon1* and *M_Expect1*, whereas *M_EmoN1* on *M_EpathyZora1*.

Table 19

Tests of Between-Subjects Effects on experiential variables (N1 =139)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon1	*M_Bon1*	1	.081	.107	.744	.001
	M_PerEmp1	1	.561	.789	.376	.006
	M_Rel1	1	2.802	2.767	.099	.021
	M_Expect1	1	3.618	3.927	.050	.029
	M_DesReal1	1	.011	.015	.901	.000
M_EmoP1	*M_Bon1*	1	3.282	4.337	.039	.032
	M_PerEmp1	1	.352	.495	.483	.004
	M_Rel1	1	.604	.596	.441	.004
	M_Expect1	1	4.145	4.498	.036	.033
	M_DesReal1	1	1.039	1.464	.228	.011
M_EmoN1	*M_Bon1*	1	.013	.017	.895	.000
	M_PerEmp1	1	3.348	4.708	.032	.034
	M_Rel1	1	.756	.747	.389	.006

	M_Expect1	1	.296	.322	.572	.002
	M_DesReal1	1	.341	.481	.489	.004
MusicEmpathy	*M_Bon1*	1	.041	.055	.815	.000
	M_PerEmp1	1	.437	.615	.434	.005
	M_Rel1	1	.800	.790	.376	.006
	M_Expect1	1	.060	.065	.799	.000
	M_DesReal1	1	.297	.418	.519	.003
TalkEmpathy	*M_Bon1*	1	.000	.001	.980	.000
	M_PerEmp1	1	.106	.149	.700	.001
	M_Rel1	1	.291	.288	.593	.002
	M_Expect1	1	.109	.118	.731	.001
	M_DesReal1	1	.091	.129	.721	.001
*MusicEmpathy * TalkEmpathy*	*M_Bon1*	1	5.464	7.220	.008	.052
	M_PerEmp1	1	4.915	6.913	.010	.050
	M_Rel1	1	3.650	3.605	.060	.027
	M_Expect1	1	.693	.752	.387	.006
	M_DesReal1	1	2.713	3.824	.053	.028

Subsequently, we checked the simple main effect of *MusicEmpathy* and *TalkEmpathy* on DVs through two-way Univariate Tests. The p value had been adjusted by Bonferroni method for multiple comparations. Taking music empathy group (*MusicEmpathyY*) into our univariate analysis, it was found that the difference on *M_PerEmp1* ($F_{(5,\ 135)}$ = 4.31, p = .040; $\eta 2$ =.032, *MD* $_{(TalkEmpathyN\ -\ TalkEmpathyY)}$ = -.43) between groups with and without talking empathy were statistically significant. Also, in the talking empathy group (*TalkEmpathyY*), music empathy significantly influenced *M_Bon1* ($F_{(5,\ 135)}$ = 4.21, p = .042; $\eta 2$ =.03, *MD* $_{(MusicEmpathyN\ -\ MusicEmpathyY)}$ = -.43) and *M_PerEmp1* ($F_{(5,\ 135)}$ = 5.76, p = .018; $\eta 2$ =.042, *MD* $_{(MusicEmpathyN\ -\ MusicEmpathyY)}$ = -.49). The significant differences were compiled in Table 20.

Table 20

Significant results of Univariate Test (N1 =139)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2
M_Bon1	*TalkEmpathyY*	*MusicEmpathyN*	*MusicEmpathyY*	-.433*	.042	4.208	.031
M_PerEmp1	*TalkEmpathyY*	*MusicEmpathyN*	*MusicEmpathyY*	-.491*	.018	5.761	.042
	MusicEmpathyY	*TalkEmpathyN*	*TalkEmpathyY*	-.433*	.040	4.314	.032

[b] Adjustment for multiple comparisons: Bonferroni

Summarizing this subsection, the interaction between *MusicEmpathy* and *TalkEmpathy* did not have statistically significant effects on the combined scores of DVs, but on *M_Bon1* and *M_PerEmp1*. In other words, the effect of *MusicEmpathy* on participants' bonding with Zora (*M_Bon1*) and perception on Zora empathy (*M_PerEmp1*) varied from whether were exposed to *TalkEmpathy* or not. There was no main effect of *MusicEmpathy* nor *TalkEmpathy* on the combination of the experience. The single main effect analysis further revealed how Media interacted to influence the participant's perception and experience. It was found that when Zora had showed talking empathy, its music empathy made the participants feel more bonding with and perceive more empathy from it (i.e., Zora). Given Zora was ready to play music for the participant, its talking empathy was necessary otherwise it blinded the participants from perceiving Zora's empathy. The result told us that the concurrence of media could give the participants' better experience.

MANOVA on experiential variables with *N1 without MOuliers* = 130. From Table 18, we found *M_EmoP1, M_EmoN1 and M_Lon1* could have correlation with the DVs (e.g., *M_Bon1* and *M_EpathyZora1*). To avoid the noise from the participants with negative mental state, we excluded the mental state outliers and re-examined the mean difference of experiential variables across Media through MANOVA.

The hypotheses of MANOVA were tested before the analyses: the Pearson correlation analyses showed no multicollinearity between any two dependent variables ($|r|<0.9$); no multivariate outliers were found by Mahalanobis distance ($p > 0.001$); Levene's test said the data did not violate the homogeneity of variance ($p > 0.05$); Box's M showed the linear correlation between random variables were negative, indicating the independence of the experiential variables ($p = .091$). (Check Appendix 1: Table 1 to 3 for details)

The interaction effect of *MusicEmpathy* and *TalkEmpathy* had no statistical significance on a combination of experiential variables (F=1.55, p=.18, $Wilks' \Lambda$=.94, *partial* η^2=.061), nor did the main effect of *MusicEmpathy* (F=.59, p=.71, $Wilks' \Lambda$=.98, η^2=0.024) and *TalkEmpathy* (F=.20, p=.96, $Wilks' \Lambda$=.99, η^2=0.008) respectively. Nonetheless, we scrutinized if there was any effects on the separateexperiential variable. The significant effects of *MusicEmpathy * TalkEmpathy* on *M_Bon1* ($F_{(1,130)}$ = 5.61, p = .019, η^2=0.044) and *M_PerEmp1* ($F_{(1,130)}$ = 4.94, p = .028. η^2=0.039 were found on Tests of Between-Subjects Effects result. (Check Appendix 1: Table 4 and 5 for details). To know where the significance was from, we further checked the simple main effect of *TalkEmpathy* and *MusicEmpathy* through the two-way Univariate Test. However, we could not find any significance in the result. In other words, it seems that the difference, caused by the Media interaction, on *M_Bon1* and *M_PerEmp1* were errorly significant. In addition, without the mental state outliers, *M_EmoP1* and *M_EmoN1* remained to be significantly associated with *M_Bon1* and *M_PerEmp1* respectively (Check Appendix 1: Table 4 to 5 for details).

MANOVA on experiential variables with *n1* = 125. Other than the mental state outliers, we still had five experiential outliers in *N1* = 139. In this section, we excluded them along with the mental state outliers. With the neat dataset *n1* = 125, we re-ran MANOVA analyses.

With *n1*= 125, no multicollinearity between any two dependent variables was found, nor was multivariate outliers. The covariance matrices of the dependent variables are equal across groups (p =.064). However, unequal variances of the *M_Bon1* (p = .023) across groups was found in Levene's test, which could increase a Type I error. Because *bonding* is the most important variable in our study, we would not exclude it but acceptted a higher alpha level of the differential significance (e.g., a bit higher than .05; Check Appendix 2: Table 1 to 3 for details)

Consistent with the analyses results with *N1* = 139 and *N1 without MOuliers* = 130 , neither interaction effect of *MusicEmpathy* * *TalkEmpathy* ($F_{(1,125)}$ = 1.96, p = .096, η^2=.078) nor main effect of *Media* (*MusicEmpathy* : $F_{(1,125)}$ = .63, p = .68, η^2=.027; *TalkEmpathy* : $F_{(1,125)}$ = .41, p = .84, η^2=.018)were found on experiencing Zora (i.e., a combination of experiential variables). *M_EmoP1* and *M_EmoN1* also had the significant moderation effect of *Media* on *M_Bon1* ($F_{(1,125)}$ = 3.17, p = .010, η^2=.12) and *M_PerEmp1* ($F_{(1,125)}$ = 2.53, p = .033, η^2=.10) respectively. We further checked the simple main effect of *Media* through two-way Univariate Test to explain the differential significance of *MusicEmpathy* * *TalkEmpathy* on *M_Bon1* and *M_PerEmp1*. (Check Appendix 2: Table 4 to 5 for details)

In the analyses result, we found that, in group who were not exposed to music empathy, *TalkEmpathy* influenced their rating to *M_Bon1* ($F_{(1,125)}$ = 6.70, p = .011; η^2 =.054, $MD_{(TalkEmpathyN\ -\ TalkEmpathyY)}$ = .52) and *M_PerEmp1* ($F_{(1,125)}$ = 4.91, p = .029; η^2 =.040, $MD_{(MusicEmpathyN\ -\ MusicEmpathyY)}$ = .43). Without talking empathy, those were exposed to music empathy (*MusicEmpathyY*) from Zora felt less bonding with Zora by .43 mean difference ($F_{(1,125)}$ = 4.50, p = .036; η^2 =.037, $MD_{(MusicEmpathyN\ -\ MusicEmpathyY)}$ = .43) and less relevant to Zora ($F_{(1,125)}$ = 5.34, p = .023; η^2 =.043, $MD_{(MusicEmpathyN\ -\ MusicEmpathyY)}$ =.54). (Check Appendix 2: Table 6 for details).

Summary. With three datasets - *N1* = 139 , *N1 without MOuliers* = 130 and *n1* = 125, we got some consistent findings:

1) In three datasets, *M_EmoP1* moderated the effect of *Media* (i.e., *MusicEmpathy* and *TalkEmpathy*) on *M_Bon1* in a positive direction, whereas *M_EmoN1* on *M_PerEmp1* in a negative way. With all participants in the first session (*N1* = 139), *M_EmoP1* was found positvely moderated the impact of *Media* on *M_Expect1*.
2) Only the co-occurrence of music and talking empathy resulted in significant difference on *M_Bon1* and *M_PerEmp1* (only interaction effects make differences).

3) We could see the crossover interaction effect of the participant experienced the co-existence and co-absence of *Media* had a better experience than those who were exposed to a sole form of empathy in terms of the bonding, perceived empathy, relevance, positive valence and design realism (Table 21 and Figure 3).

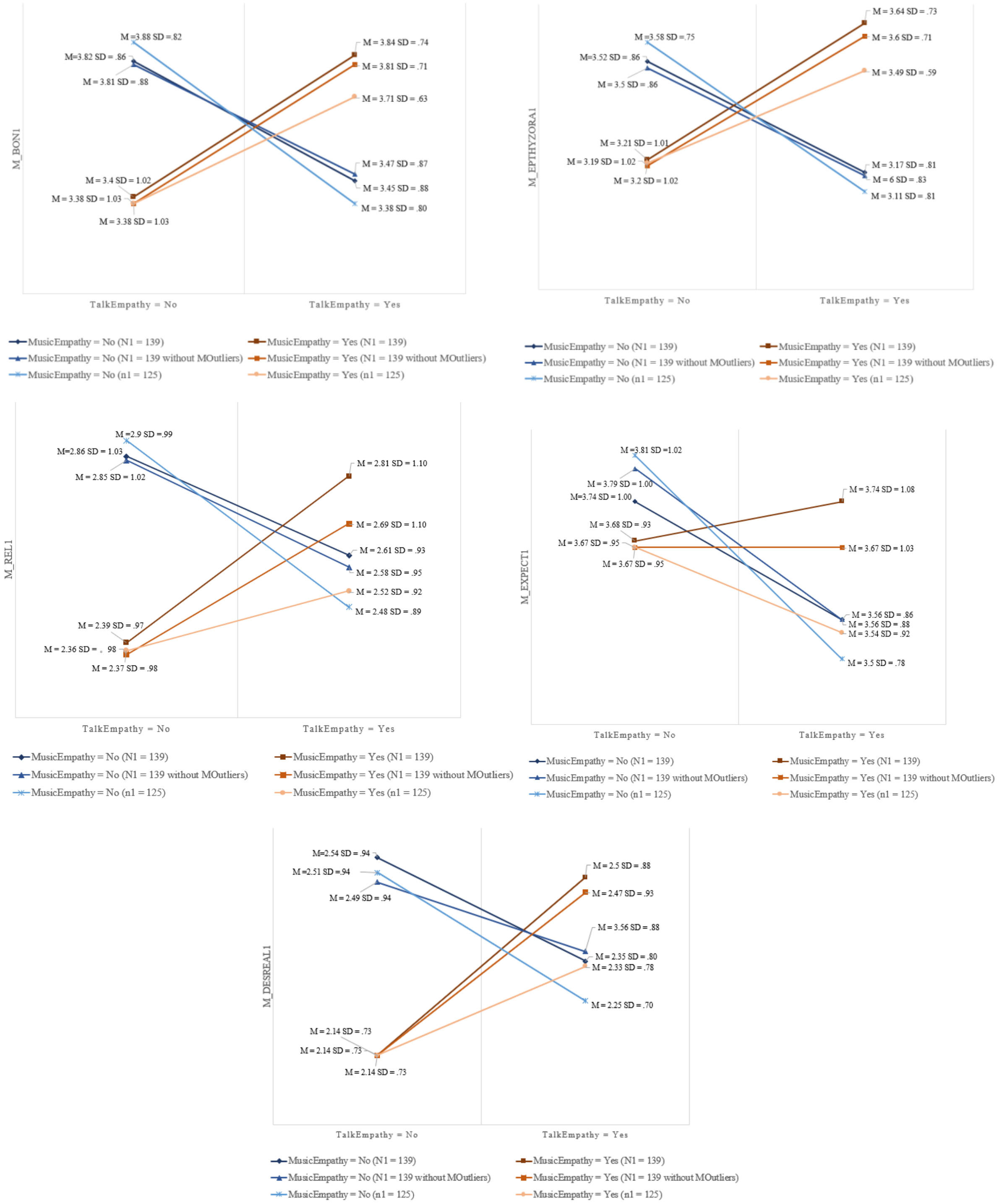


Figure 3. Crossover effect of *MusicEmpathy* and *TalkEmpathy* (data from session 1)

Table 21

Mean difference among MusicEmpathy and TalkEmapthy (data from Session 1)

	MusicEmpathy	*TalkEmapthy*	*N1* = 139		*N1* = 139 without Moutliers		*n1* = 125	
			Mean	*N*	*Mean*	*N*	*Mean*	*N*
M_Bon1	No	No	3.82 (*SD* = .86)	37	3.81 (*SD* = .88)	35	3.88 (*SD* = .82)	34
		Yes	3.45 (*SD* = .88)	36	3.47 (*SD* = .87)	34	3.38 (*SD* = .80)	32
	Yes	No	3.40 (SD = 1.02)	33	3.38 (SD = 1.03)	32	3.38 (SD = 1.03)	32
		Yes	3.84 (*SD* = .74)	33	3.81 (*SD* = .71)	29	3.71 (*SD* = .63)	27
M_PerEmp1	No	No	3.52 (*SD* = .86)	37	3.50 (*SD* = .86)	35	3.58 (*SD* = .75)	34
		Yes	3.17 (*SD* = .81)	36	3.16 (*SD* = .83)	34	3.11 (*SD* = .81)	32
	Yes	No	3.21 (*SD* = 1.01)	33	3.19 (*SD* = 1.02)	32	3.20 (*SD* = 1.02)	32
		Yes	3.64 (*SD* = .73)	33	3.60 (*SD* = .71)	29	3.49 (*SD* = .59)	27
M_Rel1	No	No	2.86 (*SD* = 1.03)	37	2.85 (*SD* = 1.02)	35	2.90 (*SD* = .99)	34
		Yes	2.61 (*SD* = .93)	36	2.58 (*SD* = .95)	34	2.48 (*SD* = .89)	32
	Yes	No	2.39 (*SD* = .97)	33	2.36 (*SD* = .98)	32	2.37 (*SD* = .98)	32
		Yes	2.81 (*SD* = 1.10)	33	2.69 (*SD* = 1.10)	29	2.52 (*SD* = .92)	27
M_Expect1	No	No	3.74 (*SD* = 1.00)	37	3.79 (*SD* = 1.00)	35	3.81 (*SD* = 1.02)	34
		Yes	3.56 (*SD* = .86)	36	3.56 (*SD* = .88)	34	3.50 (*SD* = .78)	32
	Yes	No	3.68 (*SD* = .93)	33	3.67 (*SD* = .95)	32	3.67 (*SD* = .95)	32
		Yes	3.74 (*SD* = 1.08)	33	3.67 (*SD* = 1.03)	29	3.54 (*SD* = .92)	27
M_DesReal1	No	No	2.54 (*SD* = .94)	37	2.49 (*SD* = .94)	35	2.51 (*SD* = .94)	34
		Yes	2.33 (*SD* = .78)	36	2.35 (*SD* = .80)	34	2.25 (*SD* = .70)	32
	Yes	No	2.14 (*SD* = .73)	33	2.14 (*SD* = .73)	32	2.14 (*SD* = .73)	32
		Yes	2.50 (*SD* = .88)	33	2.47 (*SD* = .93)	29	2.32 (*SD* = .76)	27

Session 2

MANOVA on experiential variables with session 2 datasets. Same as the analyses with the session 1 data, in this subsection, analyses of the effect of **IVs** (*MusicEmpathy* and *TalkEmpathy*) on five **DVs** (*M_Bon2*, *M_PerEmp2*, *M_Rel2*, *M_Expect2*, *M_DesReal2*) with *M_Lon2*, *M_EmoP2* and *M_EmoN2* as covariate were conducted using two-way MACONOVAs. Data of *N2* = 129 with and without mental state outliers (MOutliers) and *n2* = 119 were entered into the MANOVAs respectively. Multivariate outliers (participants 44 and 138) were determined by calculating the Mahalanobis distance on all three datasets (*N2* = 129, *N2 without MOutliers* = 122, *n2* = 119). No more multivariate outlier was found after removing participants 44 and 138. Thus, these two participants were excluded in the MANOVA measures, and we updated the size of the samplings as: *N2* = 127, *N2 without MOutliers* = 120 and *n2* = 117. These three datasets met the data assumptions for homoscedasticity, absence of multicollinearity and variables independence and randomness. Check the tables in Appendix 3, 4 and 5 for the results of the assumption testing.

The results of MANOVAs showed that, under the significant controlling of *M_EmoP2* and *M_EmoN2*, there was no significant multivariate effect between *TalkEmpathy* and *MusicEmpahty* on the combination of the dependent variables, nor was the main effect of *MusicEmpahty*. Yet, significant main effects were found for *TalkEmpathy* in the three datasets (*N2* = 127: F=2.79, p=.02, *Wilks' Λ*= .89, η^2= .11; *N2 without MOutliers* = 120: F= 2.83, p= .02, *Wilks' Λ*= .89, η^2= .12; *n2* = 117: F= 2.72, p= .02, *Wilks' Λ*= .89, η^2= .11). Consequently, we checked one-way Univariate Tests results of *TalkEmpathy*. However, no significance difference across groups (*TalkEmpathyY* vs. *TalkEmpathyN*) was found (Table 22).

Table 22

Univariate Tests of TalkEmpathy (data from session 2)

Dependent Variable	TalkEmpathy	*N2* = 129			*N2* =129 without MOutliers			*n2* = 119		
		p	*F*	*η2*	*p*	*F*	*η2*	*p*	*F*	*η2*
M_Bon2	Y vs. N	.215	1.553	.013	.232	1.444	.013	.141	2.199	.020
M_PerEmp2	Y vs. N	.207	1.609	.013	.206	1.618	.014	.101	2.743	.024
M_Rel2	Y vs. N	.105	2.674	.022	.148	2.118	.018	.098	2.790	.025
M_Expect2	Y vs. N	.058	3.678	.030	.056	3.715	.032	.089	2.948	.026
M_DesReal2	Y vs. N	.657	.198	.002	.914	.012	.000	.795	.068	.001

Though there were no interaction effects between *TalkEmpathy* and *MusicEmpathy* on the experience, the tests of Between-subjects effects showed that their interaction had impact on *M_Bon2* using datasets *N2* = 127 (F= 4.00, p < .05, η^2= .03) and *N2 without MOutliers* = 120 (F= 4.73, p= .03, η^2= .04). However, when we checked the two-way Univariate Test results, the statistically significant difference caused by the simple main effects was found not only on *M_Bon2* but also on *M_Rel2* and *M_Expect2* (Table 23). We could see in Table 23 with three datasets, those who were exposed to music empathy felt more bonding with Zora if it as well performed talking empathy. Among the participants who were within *N2* = 127 and restricted in the group *MusicEmpathyY*, there was statistically significant difference of their rating to *M_Rel2* between receiving talking empathy or not ($N2$ = 127: F= 4.04, p= .03, η^2= .03), indicating that if talking empathy induced more feeling of relevance in the music group. In addition, with *N2 without MOutliers* = 120 and *n2* = 117, it was found the participants had more positive valence if they interacted with Zora only played music, which reflected on the significant difference on *M_Expect2* ($N2$ = 127: MD = .49, F= 4.00, p <.05, η^2= .03; $n2$ = 117: F= .52, MD = .52 p= .04, η^2=.04).

Table 23

Significant results of Univariate Test (data from session 2)

DVs	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2	N
M_Bon2	*MusicEmpathyY*	*TalkEmpathyN*	*TalkEmpathyY*	-.511	.025	5.134	.041	127
				-.538	.020	5.543	.047	120
				-.489	.029	4.874	.042	117
M_Rel2	*MusicEmpathyY*	*TalkEmpathyN*	*TalkEmpathyY*	-.485	.047	4.044	.033	127
M_Expect2	*MusicEmpathyY*	*TalkEmpathyN*	*TalkEmpathyY*	.490	.048	4.005	.032	127
				.518	.038	4.434	.039	117

[b] Adjustment for multiple comparisons: Bonferroni

Summary. With three datasets – *N2* = 127, *N2 without MOutliers)* = 120 and *n2* = 117, we got some interesting findings:

1) Both *M_EmoP1* and *M_EmoN1* played a significant controlling effect of Media on DVs. *M_EmoP1* moderated the effect of *MusicEmpathy* and *TalkEmpathy* on the experiential variables in terms of *M_Bon2* (with *N2* = 127, *N2 without MOutliers* = 120 and *n2* = 117), *M_PerEmp2* (with *N2 without MOutliers)* = 120 and *n2* = 117), *M_Rel2* (with *N2* = 127, *N2 without MOutliers* = 120 and *n2* = 117) and *M_DesReal2* (*n2* = 117); *M_EmoN2* influenced the process of *MusicEmpathy* and *TalkEmpathy* to *M_PerEmp2* (*N2* = 127, *n2* = 117) and *M_Rel2* (*N2* = 127, *N2 without MOutliers* = 120).

2) Within participants who were exposed to the music empathy, those who as well received talking empathy from Zora tended to have a stronger bonding with Zora. This tendency appeared with three datasets. In addition among the music empathy group, the mean of *M_Expect2* on *TalkEmpathyY was lower than that* on*TalkEmpathyN,* both with the datasets *N2* = 127 and *n2* = 117. It suggested that talking Zora could let the participants down. Also, with *N2* = 127, it showed that participants felt more relevance to Zora when it showed talking empathy.

3) We found the crossover interaction effect that those were exposed to/hidden from two media had a stronger bonding with Zora. Without talking empathy, the absence of music empathy gained higher scores of the perceived empathy (*M_PerEmp2*. The negative impact of music empathy on *M_PerEmp2* could be compensated with talking empathy, thus reached the almost equal effect of the co-occurrence of music and talking empathy (check second line chart in Figure 4). *M_Rel2* demonstrated the similar tendency as *M_PerEmp2,* except that talking Zora without music empathy was scored the most relevant to the participants. The fifth line chart in Figure 3 showed that music empathy had a negative impact on the perception of design realism (*M_DesReal2*). The most interesting finding was reflected in forth line chart in Figure 3, which said if Zora not trying to perform empathy, people felt less disappointment on Zora. . However, it was just the tendency of how people experienced Zora through different media and lacked empirical evidence in our study.

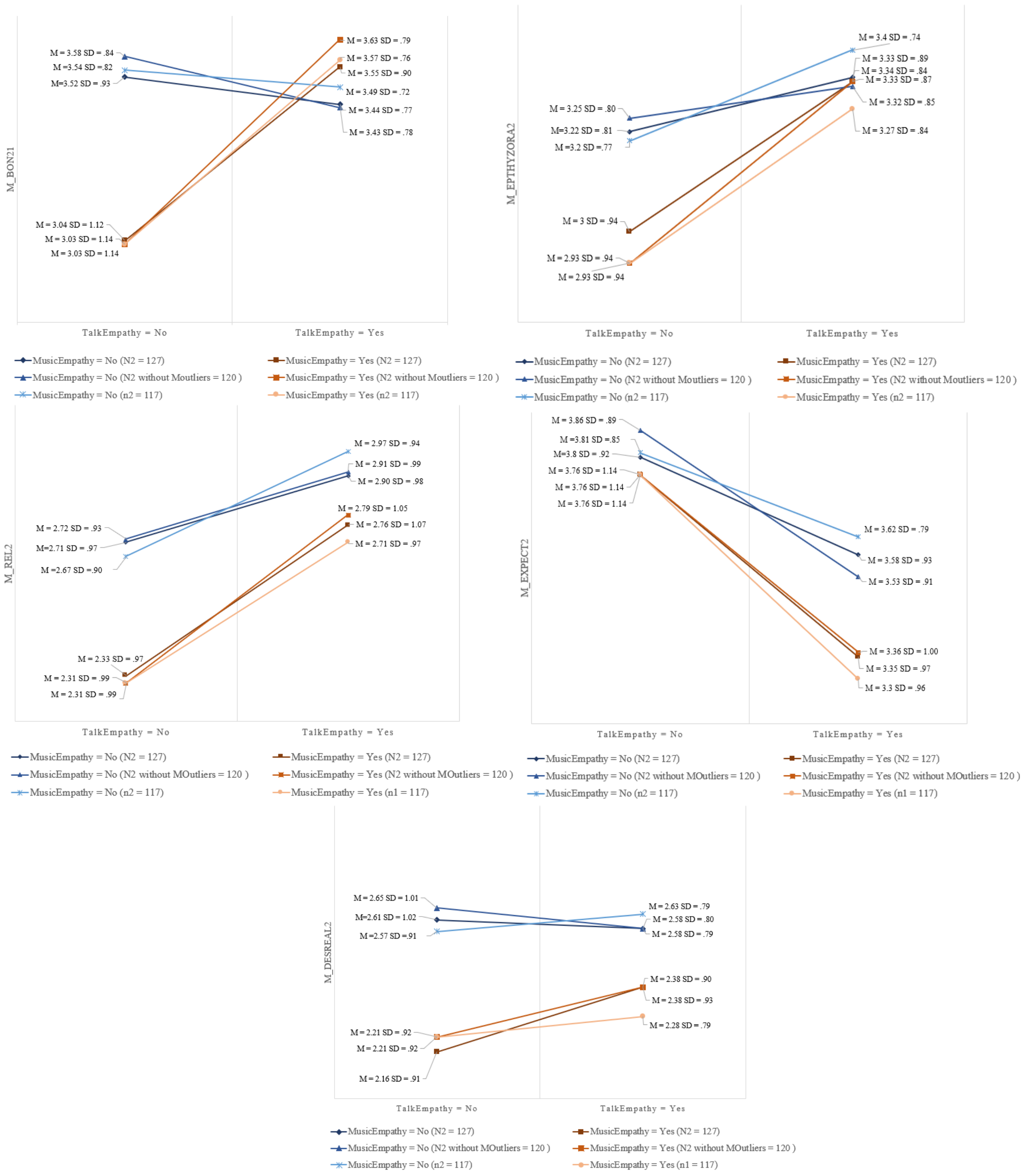


Figure 4. Crossover effect of *MusicEmpathy* and *TalkEmpathy* (data from session 2)

Table 24

Mean difference among MusicEmpathy and TalkEmpathy (data from session 2)

	MusicEmpathy	*TalkEmpathy*	*N2* = 127		*N2* $_{(MOutliers\ excluded)}$ = 120		*n2* = 117	
			Mean	*N*	Mean	*N*	Mean	*N*
M_Bon2	No	No	3.52 (*SD* = .93)	35	3.58 (*SD* = .84)	33	3.54 (*SD* = .82)	32
		Yes	3.44 (*SD* = .77)	31	3.43 (*SD* = .78)	30	3.49 (*SD* = .72)	29
	Yes	No	3.04 (*SD* = 1.12)	31	3.03 (*SD* = 1.14)	29	3.03 (*SD* = 1.14)	29

		Yes	3.55 (*SD* = .90)	30	3.63 (*SD* = .79)	28	3.57 (*SD* = .76)	27
M_PerEmp2	No	No	3.22 (*SD* = .81)	35	3.25 (*SD* = .80)	33	3.20 (*SD* = .77)	32
		Yes	3.34 (*SD* = .84)	31	3.32 (*SD* = .85)	30	3.40 (*SD* = .74)	29
	Yes	No	3.00 (*SD* = .94)	31	2.93 (*SD* = .94)	29	2.93 (*SD* = .94)	29
		Yes	3.33 (*SD* = .87)	30	3.33 (*SD* = .89)	28	3.27 (*SD* = .84)	27
M_Rel2	No	No	2.71 (*SD* = .97)	35	2.72 (*SD* = .93)	33	2.67 (*SD* = .90)	32
		Yes	2.90 (*SD* = .98)	31	2.91 (*SD* = .99)	30	2.97 (*SD* = .94)	29
	Yes	No	2.33 (*SD* = .97)	31	2.31 (*SD* = .99)	29	2.31 (*SD* = .99)	29
		Yes	2.76 (*SD* = 1.07)	30	2.79 (*SD* = 1.05)	28	2.71 (*SD* = .97)	27
M_Expect2	No	No	3.80 (*SD* = .92)	35	3.86 (*SD* = .89)	33	3.81 (*SD* = .85)	32
		Yes	3.58 (*SD* = .93)	31	3.53 (*SD* = .91)	30	3.62 (*SD* = .79)	29
	Yes	No	3.76 (*SD* = 1.10)	31	3.76 (*SD* = 1.14)	29	3.76 (*SD* = 1.14)	29
		Yes	3.35 (*SD* = .97)	30	3.36 (*SD* = 1.00)	28	3.30 (*SD* = .96)	27
M_DesReal2	No	No	2.61 (*SD* = 1.02)	35	2.65 (*SD* = 1.01)	33	2.57 (*SD* = .91)	32
		Yes	2.58 (*SD* = .79)	31	2.58 (*SD* = .80)	30	2.63 (*SD* = .78)	29
	Yes	No	2.16 (*SD* = .91)	31	2.21 (*SD* = .92)	29	2.21 (*SD* = .92)	29
		Yes	2.38 (*SD* = .90)	30	2.38 (*SD* = .93)	28	2.28 (*SD* = .79)	27

Session 3

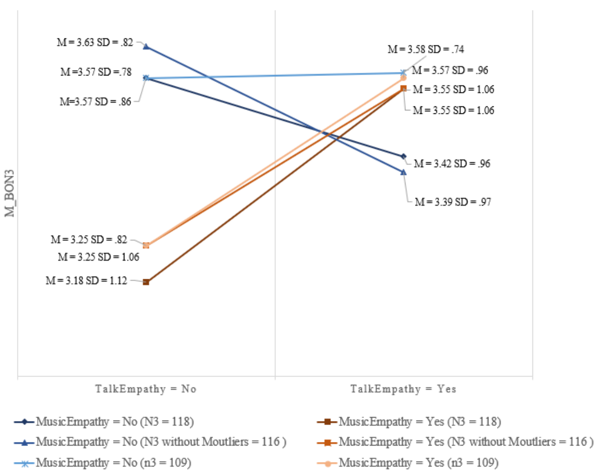
MANOVA on experiential variables with session 3 datasets. Same analyses conducted with session 3 datasets (*N3* = 121, *N3 without MOutliers* = 118, *n3* = 111). Participants 44, 59 and 138 were the multivariate outliers, and thus excluded from the datasets. It should be noted that participant 56 was also univariate mental state outliers and had been already excluded in *N3 without MOutliers* = 118 and *n3* = 111. As a result, we updated the sampling size for the datasets: *N3* = 118, *N3 without MOutliers* = 116, *n3* = 109.

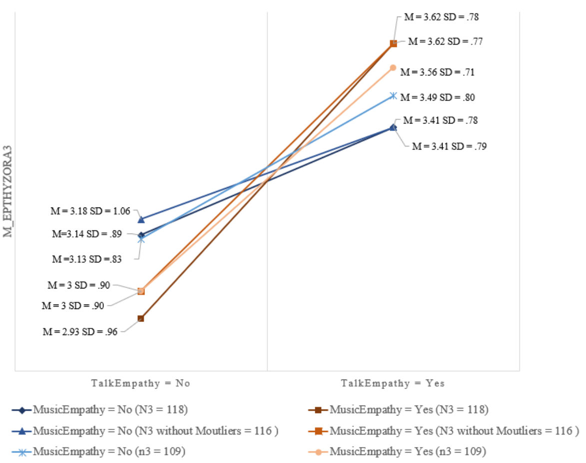
The analyses displayed the consistent results with three datasets of session 3. Firstly, there were no interaction effects between *MusicEmpathy* and *TalkEmpathy* on the combination of the experiential valiables, nor was the main effects of *MusicEmpathy*. Check Appendix 6-8 for the results. Regarding to the main effects of *TalkEmpathy*, results showed statistically significant difference across groups (*N3* = 121: F=8.19, p=.000, *Wilks' Λ*= .72, η^2= .28; *N3 without MOutliers* = 118: F= 7.75, p= .000, *Wilks' Λ*= .73, η^2= .27; *n3* = 111: F= 7.11, p= .000, *Wilks' Λ*= .73, η^2= .27). These significances were reflected in terms of *M_PerEmp3* and *M_Expect3* (Check the Appendix 6-8, Table 6 for Univariate Tests for

TalkEmpathy). After we took into the analyses the participants who were exposed to music empathy, they were found, with talking empathy, perceived more empathy from Zora, however, felt more disappointment (Check the Appendix 6-8, Table 7).

Summary. With three datasets of session 3 observations, we found

1) Those who received talking empathy had higher score in *M_PerEmp3* and lower score in *M_Expect3*. Especially, within *MusicEmpathyY* samplings, talking empathy played positive effect on *M_PerEmp3* and *M_Expect3*.
2) Different from the results with datasets of other two sessions, in the third session the participants perceived talking robot more empathic, regardless of its playing music or not (second line chart in Figure 4). The results interested us that the more actions Zora did the less positive valence on Zora the participants had, which was also revealed with datasets from first and second sessions.
3) *Media* still displayed a crossover interaction effect on *Bonding* and *Relevance* (i.e., *M_Bon3* and *M_Rel3*), showing that the co-occurrence/co-absence of *Media* brought better experience.
4) Same as the result with datasets from session 1 and 2, the participants rated the below 3 (3 represented a little disagree) on Design Realism no matter how Zora performed. It indicated that they thought Zora looked different from human.





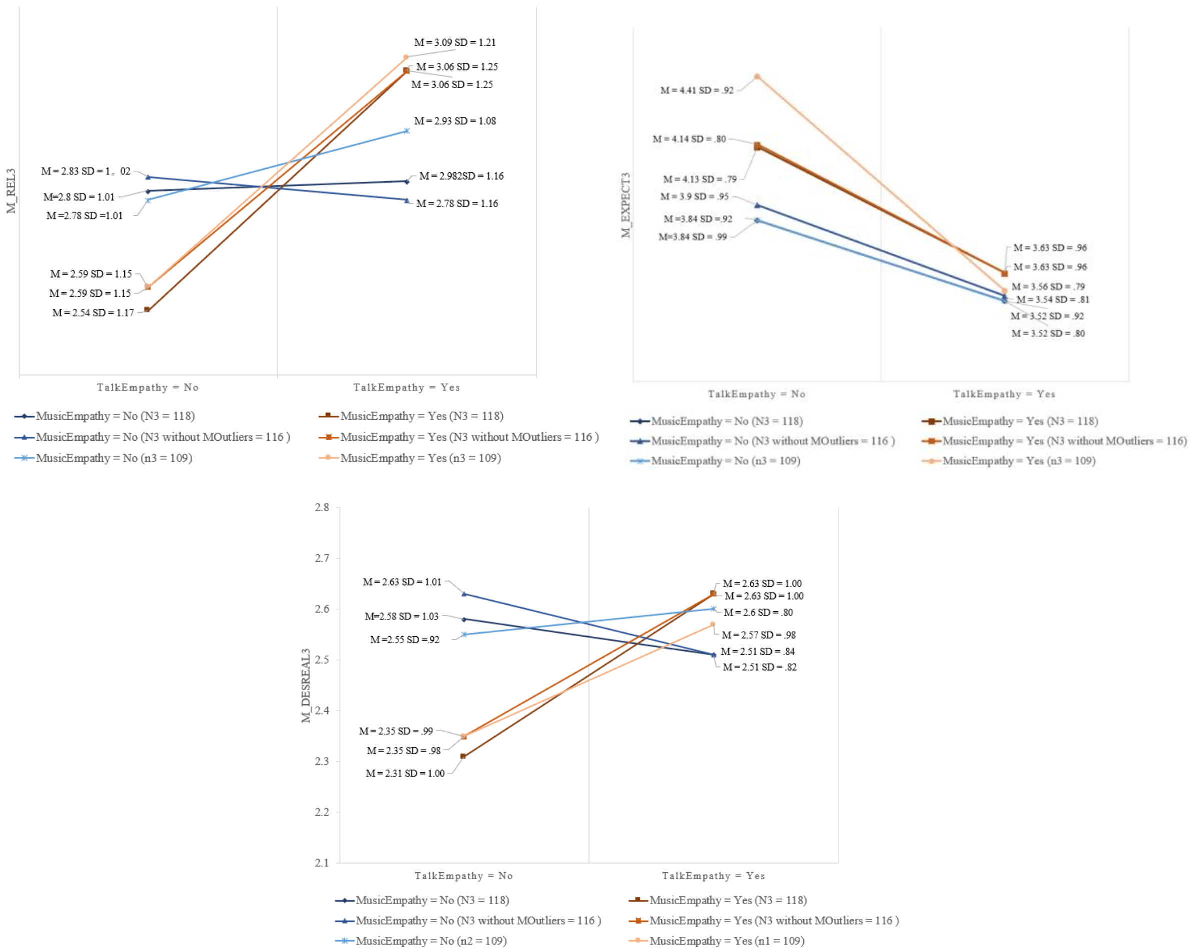


Figure 5. Crossover effect of *MusicEmpathy* and *TalkEmapthy* (data from session 3)

Table 25

Mean difference among MusicEmpathy and TalkEmapthy (data from session 3)

	MusicEmpathy	*TalkEmpathy*	*N3* = 118 Mean	*N*	*N3*$_{(MOutliers\ excluded)}$ = 116 Mean	*N*	n3 =109 Mean	*N*
M_Bon3	No	No	3.57 (*SD* = .86)	31	3.63 (*SD* = .82)	30	3.57 (*SD* = .78)	29
		Yes	3.42 (*SD* = .96)	27	3.39 (*SD* = .97)	26	3.58 (*SD* = .74)	24
	Yes	No	3.18 (*SD* = 1.12)	30	3.25 (*SD* = .82)	29	3.25 (*SD* = 1.06)	29
		Yes	3.55 (*SD* = 1.06)	30	3.55 (*SD* = 1.06)	30	3.57 (*SD* = .96)	27
M_PerEmp3	No	No	3.14 (*SD* = .89)	31	3.18 (*SD* = 1.06)	30	3.13 (*SD* = .83)	29
		Yes	3.41 (*SD* = .78)	27	3.41 (*SD* = .79)	26	3.49 (*SD* = .80)	24
	Yes	No	2.93 (*SD* = .96)	30	3.00 (*SD* = .90)	29	3.00 (*SD* = .90)	29
		Yes	3.62 (*SD* = .78)	30	3.62 (*SD* = .77)	30	3.56 (*SD* = .71)	27
M_Rel3	No	No	2.80 (*SD* = 1.01)	31	2.83 (*SD* = 1.02)	30	2.78 (*SD* = 1.01)	29
		Yes	2.82 (*SD* = 1.16)	27	2.78 (*SD* = 1.16)	26	2.93 (*SD* = 1.08)	24
	Yes	No	2.54 (*SD* = 1.17)	30	2.59 (*SD* = 1.15)	29	2.59 (*SD* = 1.15)	29

		Yes	3.06 (*SD* = 1.25)	30	3.06 (*SD* = 1.25)	30	3.09 (*SD* = 1.21)	27
M_Expect3	No	No	3.84 (*SD* = .99)	31	3.90 (*SD* = .95)	30	3.84 (*SD* = .92)	29
		Yes	3.52 (*SD* = .80)	27	3.54 (*SD* = .81)	26	3.52 (*SD* = .92)	24
	Yes	No	4.13 (*SD* = .79)	30	4.14 (*SD* = .80)	29	4.14 (*SD* = .92)	29
		Yes	3.63 (*SD* = .96)	30	3.63 (*SD* = .96)	30	3.56 (*SD* = .79)	27
M_DesReal3	No	No	2.58 (*SD* = 1.03)	31	2.63 (*SD* = 1.01)	30	2.55 (*SD* = .92)	29
		Yes	2.51 (*SD* = .82)	27	2.51 (*SD* = .84)	26	2.60 (*SD* = .80)	24
	Yes	No	2.31 (*SD* = 1.00)	30	2.35 (*SD* = .99)	29	2.35 (*SD* = .98)	29
		Yes	2.63 (*SD* = 1.00)	30	2.63 (*SD* = 1.00)	30	2.57 (*SD* = .98)	27

Effects of Media on experiential variables over Session

In the previous sections, we assessed whether Zora empathy through different *Media* could induce diverse experiences on the interaction. We measured the experience three times. In this section, we examined the time effects on the experience. Due to repeated measurement, there should be a certain degree of correlation between the measurements of individual, which might violates the requirement of data independence in the analysis of variance. Thus, the premise of the repeated analyses of variance is that within-subjects variables at each time point should satisfy the sphericity assumption. Mauchly method is usually used to test whether the spherical assumption is satisfied. If the test result is $p > 0.05$, the assumption of data is satisfied; if $p < 0.05$, it is not satisfied. When the data satisfies the spherical assumption, the one-way analysis of variance can be directly carried out; if not, the results of the multivariate analysis of variance prevails.

In our study, we used *MusicEmpathy* (yes-no), *TalkEmpathy* (yes-no) and *Session* (first – second - third) as independent factors and then we have a 2*2*3 Repeated MANOVA measure on Bonding (*M_Bon*), Perceived Empathy (*M_PerEmp*), Relevance (*M_Rel*), Valence (*M_Expect*) and Design Realism (*M_DesReal*), with Loneliness (*M_Lon*), Positive emotion (*M_EmoP*) and Negative emotion (*M_EmoN*) as covariate. To more precisely, it was a three-factor mixed experiment with repeated measures of five factors.

We firstly checked the result of the intra-group comparison. In our case, most of the spherical test results $p<0.05$ (Table 26), which indicated that the data except for *M_Expect* did not meet the spherical assumption. Thus, the results of multivariate analysis of variance prevailed (Table 27). Moreover, we referred to the corrected one-way Univariate results where the Greenhouse-Geisser correction results were recommended (Table 28). In Table 27, both *Session* and *Session * TalkY* had $p < 0.05$, indicating that there was difference on the experiential variables over sessions. More particularly, the experience of *M_Bon* (F= 6.80, p=.004, η^2=0.07), *M_PerEmp* (F=3.43, p=.038, η^2=0.04) and *M_Rel* (F= 3.92, p=.027, η^2=0.04) changed with session and the effect of talking empathy on *M_PerEmp* (F= 5.49, p=.006, η^2=0.06) varied along with session. Neither interaction effect nor main effect of music empathy on experiential variables across sessions were found.

Table 26

Mauchly's Test of Sphericity[a]

Within Subjects Effect	Measure	Mauchly's W	Approx. Chi-Square	*df*	Sig.	Epsilonb		
						Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Session	*M_Bon*	.690	33.051	2	.000	.763	.800	.500
	M_PerEmp	.926	6.835	2	.033	.931	.982	.500
	M_Rel	.851	14.310	2	.001	.871	.916	.500
	M_Expect	.991	.769	2	.681	.991	1.000	.500
	M_DesReal	.789	21.060	2	.000	.826	.867	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identity matrix.
a. Design: Intercept + MusicY + TalkY + MusicY * TalkY Within Subjects Design: Session

Table 27

Wilks' Lambda test result of Multivariate analyses (repeated measurements)

Within Subjects Effect	Value	*F*	Hypothesis df	Error df	*Sig.*	*η2*
Session	.783	4.575	10.000	352.000	.000	.115
*Session * MusicY*	.927	1.362	10.000	352.000	.197	.037
*Session * TalkY*	.878	2.366	10.000	352.000	.010	.063
*Session * MusicY * TalkY*	.952	.868	10.000	352.000	.563	.024

a. Design: Intercept + MusicY + TalkY + MusicY * TalkY
Within Subjects Design: Session

Table 27

Greenhouse-Geisser test result of Univariate analyses (repeated measurements)

Source	Measure	*df*	Mean Square	*F*	*Sig.*	*η2*
Session	*M_Bon*	1.526	2.371	6.799	.004	.070
	M_PerEmp	1.862	.967	3.432	.038	.037
	M_Rel	1.741	1.073	3.918	.027	.042
	M_Expect	1.983	.392	.921	.399	.010
	M_DesReal	1.652	.555	2.674	.083	.029
*Session * MusicY*	*M_Bon*	1.526	.127	.365	.638	.004
	M_PerEmp	1.862	.429	1.522	.222	.017
	M_Rel	1.741	.528	1.926	.155	.021
	M_Expect	1.983	.628	1.474	.232	.016
	M_DesReal	1.652	.208	1.003	.356	.011
*Session * TalkY*	*M_Bon*	1.526	.048	.139	.813	.002
	M_PerEmp	1.862	1.545	5.485	.006	.057
	M_Rel	1.741	.089	.326	.692	.004
	M_Expect	1.983	.126	.296	.742	.003
	M_DesReal	1.652	.175	.844	.413	.009
*Session * MusicY * TalkY*	*M_Bon*	1.526	.394	1.129	.314	.012
	M_PerEmp	1.862	.561	1.993	.143	.022
	M_Rel	1.741	.274	1.001	.361	.011
	M_Expect	1.983	.641	1.504	.225	.016
	M_DesReal	1.652	.179	.864	.405	.010

Next, to explore under what circumstances there was a statistically significant difference between sessions, we checked the results of multiple comparisons of within-subjects' factors. The significance had been compiled in Table 29. Alpha value had been adjusted for multiple comparisons by the Bonferroni correction method. Significant difference between sessions on *M_Bon* occurred between Session1 vs Session 2 (*MD*= .26, *p*=.000) and Session 1 vs Session 3 (*MD*= .22, *p*=.022). It indicated that the participants felt more bonding (*M_Bon*) with Zora in the first session compared to that in the second and the third sessions. Moreover, the participants perceived the most empathy (*M_PerEmp*) from Zora in the first session in comparison to the mean score in the second session (*MD*= .20, *p*=.012). There was statistical significance on Relevance (*M_Rel*) between the first and the third session (*MD*= -.20, *p*=.027) and between the second and the third session (*MD*= -.16, *p*=.009). The participants in the third session reported the highest relevance of Zora.

Table 29

Independent Pairwise Comparisons between Sessions

Measure	(I) Session	(J) Session	Mean Difference (I-J)	Std. Error	Sig.
M_Bon	1	2	.260*	.071	.000
		3	.218*	.094	.022
M_PerEmp	1	2	.196*	.076	.012
M_Rel	1	3	-.186*	.083	.027
	2	3	-.157*	.059	.009

Regarding the significant main effect of talking empathy on experiential variables across the session, it was accumulated by the difference in *M_Bon* from the no-talking empathy group between session 1 and session 2 (*MD*= 29, *p*=.008) and the difference in *M_PerEmp* from the no-talking empathy group between session 1 and session 2 (*MD*= .33, *p*=.006) as well as session 1 and session 3 (*MD*= .36, *p*=.005). From Table 18 we could see no significant main effect of music empathy was found on a repeated univariate measure (refer to Table 28). However, we could see the difference in experiential variables across sessions fell to the music empathy group. Please check Table 30 for details.

Table 30

*Independent Pairwise Comparisons among TalkEmpathy*Session and MusicEmpathy*Session*

Measure	Restricted Group	(I) Session	(J) Session	Mean Difference (I-J)	*Std. Error*	*Sig.*
M_Bon	*MusicEmpathyY*	1	2	.324*	.100	.005
	TalkEmpathyN	1	2	.293*	.095	.008
M_PerEmp	*MusicEmpathyY*	1	2	.324*	.108	.010
	TalkEmpathyN	1	2	.327*	.102	.006

			3	.361*	.111	.005
M_Rel	*MusicEmpathyY*	1	3	-.299*	.117	.038
		2	3	-.287*	.083	.003
M_DesReal	*MusicEmpathyY*	2	3	-.177*	.065	.025

Summary

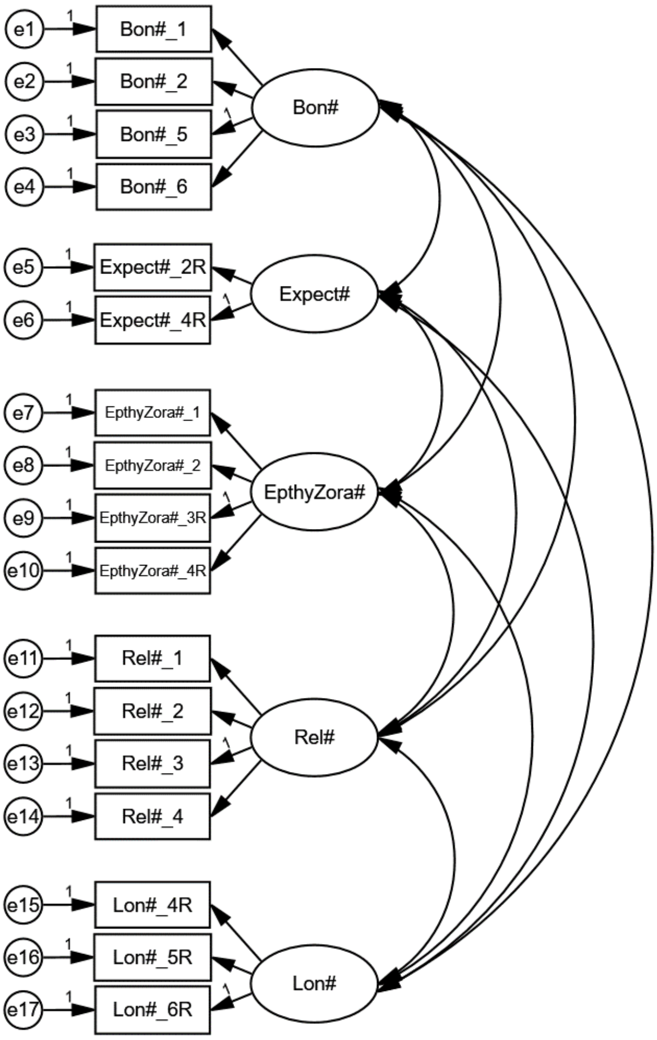
There were statistically significant differences across sessions in terms of Bonding, Perceived Empathy and Relevance. When we scrutinized where the difference came from, we found they mainly came from the non-talking empathy group. Participants reported they experienced the most bonding and Zora empathy in the first session, whereas the most feeling of relevance in the third session. Talking empathy also had an effect, mainly on Perceived Empathy, when we compared the scores for the first session with the other two sessions. The interaction effect of *Media* and the main effect of music empathy across sessions were absent in our sampling.

Multivariate Regression Analysis

Structural equation modelling (SEM) was used in this section to test and evaluate multivariate causal relationships. We ran the analyses with nine datasets. It can test direct and indirect effects on pre-assumed casual relationships and allow more than one independent variables, which outperforms Hayes Process. SEM is composed of the measurement model and the structural model. A measurement model measures the latent variables or composite variables (Smith & Hoyle, 1996; Hoyle,2011; Kline 2010), while the structural model tests all the hypothetical dependencies based on path analysis (Hoyle 1995, 2011; Kline 2016).

The estimation of sample size is the prior issue for the SEM application. However, there is no golden rules for estimation of the sample size. Some general rules recommended were, for example, a minimum sample size of 100-200 or five cases per free parameter in the model (Tabachnick and Fidell 2001; Kline 2010). Increasingly, use of model-based methods for estimation of sample size is highly recommended, with sound methods based on fit indices or power analysis of the model. Among them, Soper (2022) developed the sample size calculator for Structural Equation Models based on the work of Cohen (1988) and Westland (2010). In our study, we had 17 observed variables constructing five latent variables (check Figure 6 for the model). Given .30 anticipated effect size, .80 power level and .05 probability level, the recommended minimum sample size from the Soper's calculator was 150. It was pity that the maximum size of the nine datasets was 139 and could not reach the recommended sample size. The insufficient sample size could lead to overfitting. However, we argued our sample size was acceptable for our SEM analysis because the main path analysis in fact had less variables. The constructed variable *Lon#* containing three observed variables (*Lon#_4R*, *Lon#_5R*, *Lon#_6R*) was not included in the main path analysis, instead we explored its impact on the additional moderation analysis.

Next in this section, according to the PCA results, we firstly tested the fitness of the measurement model with each of the nine datasets; and secondly, we checked the fitness of the structural model and evaluated the psychology paths after exposure to Zora empathy; lastly, we explored whether loneliness moderated the impact of perceived Zora empathy on the expectation met level and the relevance to Zora.



 # The index of sessions

Figure 6. *Measurement Model*

Evaluation of the Measurement Model

Confirmatory Factor Analysis (CFA) was computed using AMOS to test the measurement model (Figure 6) with nine datasets respectively. As part of confirmatory factor analysis, factor loadings were assessed for each (refer to Table 31) and most of the items reached the acceptable factor loading (>.60; Awang, 2014). Only *Expect#_2R* had factor loadings greater than .50 but less than .60 in *N2*= 129 and *n2* = 119, and *PerEmp#_4R* in *n1* = 125 and *N2* = 129. Awang (2015) suggested that the newly developed items with the factor loadings exceeding .50 are acceptable. Thus, *Expect#_2R* and *PerEmp#_4R* were reserved since they were newly developed for this study and their factor loadings fluctuated between .50 and .70 above.

Next, the model-fit measures were used to assess the model's overall goodness of fit (CMIN/df, GFI, CFI, IFI, TLI, RMR, and RMSEA) and most values (CMIN/df, CFI, IFI, TLI, RMR) were within their respective common acceptance levels (Ullman, 2001; Hu and Bentler, 1998, Bentler, 1990). Regarding the Chi-Square test, the significance in *p*-values indicated our model was ill-fitting. However, Collier (2020) pointed out the chi-square test could be problematic for the large sample size and complex model and suggested a better option of the "relative chi-square" test to reduce the effect of sample size. The "relative chi-square" test is the chi-square divided by the degrees of freedom (CMIN, χ^2/df) and with a value under 3 is considered an acceptable fit (Kline, 2016). CMIN column in Table 32 suggested the measurement model fitting well the data in all nine datasets. Contrary to those acceptable indices, RMSEA of *N2* = 129 denoted an almost poor fit. The possible reason could be our measurement model fit a baseline model (i.e., a model with the worst fit) which the CFI and TLI suggested but it was far away from the perfect model that RMSEA suggested. Since only one RMSEA value exceed the adequate threshold across nine datasets, we did not modify the measurement model and would interpret the goodness of fit of the model carefully.

Finally, we concluded that the five-construct model (*M_Lon#*，*M_Bon#*, M_Expect#, M_Rel# and *M_Epthy#*) yielded the all-over acceptable fit (Table 32) for nine sets of the data across three sessions.

With the measurement model unchanged, we skipped the reliability and validity analysis of the items, as they had been discussed in ***Reliability Analysis***,***Validity Analysis*** and ***Reliability Analysis after PCA*** sections. Also, in the structural model analysis, we continued to use the construct noun representing the mean score of the items within same construct, rather than the latent variables.

Table 31

Factor Loading of Measurement Model across nine datasets

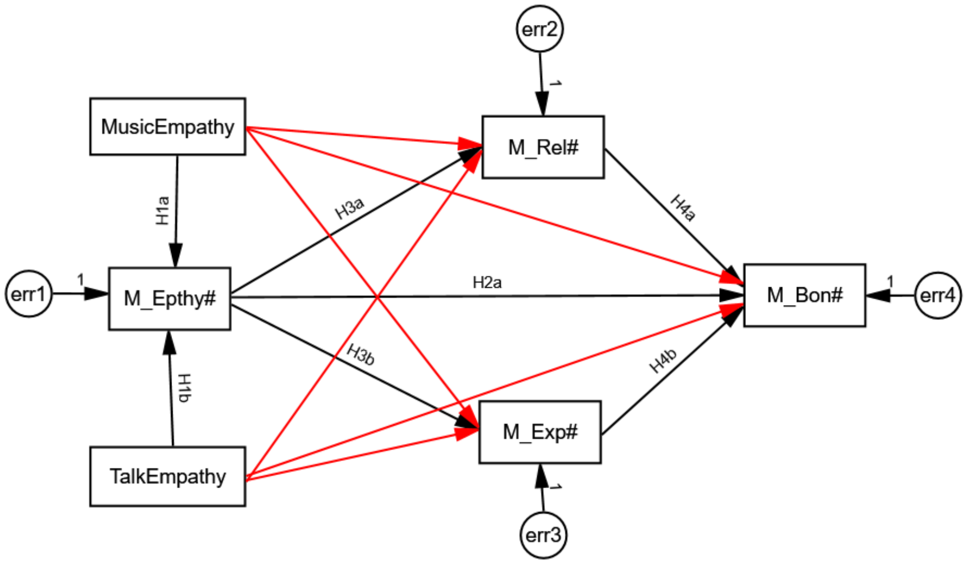
Indicators	Unobservable variables	Factor loadings: Session 1 *N1* = 139	*N1* without MOutlier = 130	*n1* = 125	Session 2 *N2* = 129	*N2* without MOutlier = 122	*n2* = 119	Session 3 *N3* = 121	*N3* without MOutlier = 117	*n3* = 112
Bon#_1	*M_Bon#*	.690	.684	.641	.876	.861	.851	.837	.814	.783
Bon#_2		.886	.884	.881	.929	.922	.916	.934	.929	.914
Bon#_5		.833	.829	.816	.868	.852	.856	.911	.900	.883
Bon#_6		.839	.833	.817	.873	.872	.862	.867	.856	.856
Expect#_2R	*M_Expect#*	.623	.612	.605	.537	.610	.563	.722	.739	.755
Expect#_4R		.985	.977	.991	1.050	1.033	1.067	1.035	.986	.959
PerEmp#_1	*M_PerEmp#*	.753	.750	.712	.896	.893	.890	.838	.813	.778
PerEmp#_2		.660	.647	.603	.739	.724	.704	.787	.751	.731
PerEmp#_3R		.707	.742	.721	.637	.711	.677	.689	.675	.781
PerEmp#_4R		.613	.605	.578	.555	.649	.611	.634	.694	.719
Rel#_1	*M_Rel#*	.934	.931	.943	.951	.949	.946	.942	.939	.936
Rel#_2		.912	.905	.897	.929	.932	.925	.939	.937	.933
Rel#_3		.811	.807	.779	.859	.851	.839	.907	.902	.895
Rel#_4		.849	.851	.850	.932	.933	.929	.944	.941	.937
Lon#_4R	M_Bon#	.822	.817	.836	.828	.827	.825	.877	.846	.835
Lon#_5R		.669	.604	.611	.857	.840	.836	.912	.888	.876
Lon#_6R		.737	.636	.618	.911	.901	.896	.894	.867	.914

Table 32

Fit Indice of Measurement Model across nine datasets

Fit Indices	Recommended Value	Source (s)	Obtained Value: Session 1 *N1* = 139	*N1* without MOutlier = 130	*n1* = 125	Session 2 *N2* = 129	*N2* without MOutlier = 122	*n2* = 119	Session 3 *N3* = 121	*N3* without MOutlier = 117	*n3* = 112
p	Insignificant	Bagozzi and Yi (1988)	.000	.000	.000	.000	.000	.000	.006	.016	.015
CMIN (χ^2/df)	< 3	Kline (2011)	1.676	1.757	1.759	2.018	1.741	1.703	1.369	1.311	1.316
GFI	>.80 adequate fit >.90 good fit	Hair et. al. (2010) Baumgartner and Homburg (1995)	.865	.855	.850	.829	.841	.841	.880	.880	.877
CFI	>.90	Bentler (1990)	.945	.932	.923	.940	.953	.951	.978	.979	.977
TLI	>.90	Bentler (1990)	.931	.915	.904	.925	.941	.939	.972	.974	.972
RMR	<.08	Hu and Bentler (1998)	.069	.071	.073	.069	.054	.054	.059	.054	.050
RMSEA	<.08 adequate fit > .10 poor fit	Hu and Bentler (1998) (MacCallum et al. 1996)	.070	.077	.078	.089	.078	.077	.055	.052	.053

Estimates of the Structural Model



 # The index of sessions

Figure 7. *Structural Model*

A structural equation model (Figure 7) generated through AMOS was used to test the relationships depicted in the hypothesized theory model. The black arrowed lines presented the relationships we hypothesized. Though the red arrowed lines were not we assumed, with them we could inspect whether the experience and perception of the participant on the empathetic Zora relied on how much they perceived empathy from Zora. In other words, whether the perceived empathy (*M_Epthy#*) was the mediator for a better experience.

A good-fitting model was accepted if the value of the $\chi2$/df less than 3, and the goodness-of-fit (GFI) indices (Hair et al., 2010), the Tucker and Lewis (1973) index (TLI) and the Confirmatory fit index (CFI) (Bentler, 1990) are ⩾ 0.90 (Hair et al., 2010). In addition, an adequate-fitting model was accepted if the value of the standardized root mean square residual (RMR) computed by AMOS< 0.05, and the root mean square error approximation (RMSEA) is between .05 and .08 (Hair et al., 2010). The fit indices for the model are shown in the last row of Table 33, Table 34 and Table 35 and they were within the acceptable range.

The squared multiple correlations revealed how much of the variance is being explained with the independent variables (refer to last-second sections of Table 33, Table 34 and Table

35). Taking *N1* = 139 as an example, the squared multiple correlation was .002 for *M_Epthy1*, this shows that .02% variance in *M_Epthy1* is accounted for by *MusicEmapthy* and *TalkEmpathy*. The small effect of *MusicEmapthy* and *TalkEmpathy* on the felt empathy from Zora (*M_Epthy#*) replicated on the analysis with otherdatasets.

The study assessed the impact of *TalkEmpathy* and *MusicEmpahty* on the felt empathy from Zora (*M_Epthy#*), that of *M_Epthy#* on the bonding with Zora (*M_Bon#)*, the relevance (*M_Rel#*) and the expectation met (*M_Expect#*), and that of *M_Rel#* and *M_Expect#* on *M_Bon#*. In *N1* = 139, the impact of *M_Epthy1* on *M_Bon1* (b= .125, *s.e.*= .082, p <.002), *M_Rel1*(b= .503, *s.e.*= .085, p <.000) and *M_Exp1*(b= .482, *s.e.*= .082, p <.000) were positive and significant, supporting H2a, H3a and H3b. The impact of *M_Rel1* on *M_Bon1* was also positive and significant (b= .457, *s.e.*= .063, p <.000), supporting H4a. Those significances were replicated in the other two datasets of the session 1 (*N1 without MOutlier* = 130 and *n1* = 125).

When inspecting the relationships among the variables with the session 2 data, we found two more significant regression impact which were *TalkEmpathy* on *M_Exp2* and *M_Expect2* on *M_Bon2* (H4b). However, the impact power of *M_Expect2* on *M_Bon2* was tiny (*N2* = 129: b= .079, *s.e.*= .076, p <.000; *N2 without MOutlier* = 122: b= .062, *s.e.*= .150, p <.000; *n2* = 119: b= .064, *s.e.*= .076, p <.000), thus H4b was accepted with small effect power. In all three datasets of session 2, the impact of *TalkEmpathy* was negative with a medium power (less than .03) on *M_Expect2.*

In the third session, *TalkEmpahty* positively explained the felt empathy from Zora (*M_Epthy3*), which supported H1a. Besides, with *N3* = 121 and *N3 without MOutlier* = 117, the impact of *TalkEmpathy* on *M_Bon3* was negatively significant tiny (*N3* = 121: b= -.142, *s.e.*= .104, p <.000; *N3 without MOutlier* = 117: b= -.133, *s.e.*= .107, p <.002).

In summary, we had an insight that, after repeated exposure to the taking empathy from Zora (*TalkEmpahy*), talking robot with empathetic response disappointed the participants (c.f. negative effect of *TalkEmpahy* on *M_Expect2* and of *TalkEmpahy* on *M_Expect3*). With time, participants recognized Zora trying to show empathy on them (c.f. positive effect of *TalkEmapthy* on *M_Expect3*) and felt less bond with Zora (c.f. negative effect of *TalkEmapthy* on *M_Bon3*). This finding is identical with what we found in ***Effects of Media on experiential variables over Session***. However, given there were two findings has found: 1) *TalkEmpahty* led to more Bonding (i.e., more *M_Bon3*) if participant perceived more empathy from Zora (i.e., more *M_Epthy3*) and 2) *TalkEmpahty* could result in more disappointment (less *M_Expect3*) and further destroy *M_Bon3*, there came a question: could perceiving Zora tried to do goodness compensate the disappointment Zora talking induced? We would inspect the media effect of *Perceived empathy* on *Bonding* in the next sub-section. Model fit indices and Hypotheses results are presented in Table 33, 34 and 35.

Table 33

Path Analysis of Structural Model with Session 1 data

	N1 = 139			*N1 without MOutlier* = 130			*n1* = 125		
	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*
Hypothesized Relationships									
H1a: *TalkEmpahty → M_Epthy1*	.011	.148	.896	.004	.154	.960	-.064	.148	.472
H1b: *MusicEmpahty → M_Epthy1*	.047	.148	.580	.030	.154	.733	-.017	.148	.848
H2a: *M_Epthy1 → M_Bon1*	.125	.082	.002	.247	.087	.004	.240	.088	.005
H3a: *M_Epthy1→ M_Rel1*	.503	.085	.000	.527	.086	.000	.466	.091	.000
H3b: *M_Epthy1 → M_Expect1*	.482	.082	.000	.508	.083	.000	.451	.088	.000
H4a: *M_Rel1 → M_Bon1*	.457	.063	.000	.469	.067	.000	.443	.070	.000
H4b: *M_Expect1 → M_Bon1*	.125	.066	.080	.100	.070	.185	.116	.072	.137
Other Relationships									
TalkEmpahty → M_Rel1	.026	.148	.719	.002	.150	.983	-.047	.150	.550
TalkEmpahty → M_Bon1	-.006	.111	.928	.013	.115	.838	.000	.119	.997
TalkEmpahty → M_Expect1	-.042	.143	.575	-.065	.144	.388	-.095	.146	.234
MusicEmpahty → M_Rel1	-.092	.149	.211	-.113	.150	.128	-.130	.150	.097
MusicEmpahty → M_Bon1	.004	.111	.945	.006	.116	.932	.009	.119	.895
MusicEmpahty → M_ Expect1	.010	.143	.898	-.017	.145	.820	-.022	.146	.779
Squared Multiple Correlation (R2):									
M_Epthy1	.002			.001			.004		
M_Rel1	.234			.262			.219		
M_ Expect1	.258			.287			.241		
M_Bon1	.462			.463			.414		

Model Fit Statistics:									
	χ^2/df = 1.59, *p* = .203, *GFI* = 992, *TLI* = 942, *CFI* = .992, *RMR* = .026, *RMSEA* = .066			χ^2/df = .726, *p* = .484, *GFI* = .996, *TLI* = 1.027, *CFI* = 1.000, *RMR* = .017, *RMSEA* = .000			χ^2/df = .350, *p* = .705, *GFI* = .998, *TLI* = 1.083, *CFI* = 1.000, *RMR* = .011, *RMSEA* = .000		

Table 34

Path Analysis of Structural Model with Session 21 data

	N2 = 129			*N2 without MOutlier* = 122			*n2* = 119		
	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*
Hypothesized Relationships									
H1a: *TalkEmpahty → M_Epthy2*	.138	.154	.113	.140	.159	.119	.168	.153	.061
H1b: *MusicEmpahty → M_Epthy2*	-.096	.154	.271	-.114	.159	.202	-.146	.153	.103
H2a: *M_Epthy2 → M_Bon2*	.310	.076	.000	.379	.075	.000	.376	.076	.000
H3a: *M_Epthy2→ M_Rel2*	.580	.082	.000	.598	.081	.000	.553	.087	.000
H3b: *M_Epthy2 → M_Expect2*	.571	.083	.000	.568	.085	.000	.519	.091	.000
H4a: *M_Rel2 → M_Bon2*	.568	.058	.000	.514	.059	.000	.515	.060	.000
H4b: *M_Expect2 → M_Bon2*	.079	.076	.000	.062	.150	.002	.064	.076	.000
Other Relationships									
TalkEmpahty → M_Rel2	.086	.143	.224	.098	.143	.171	.106	.146	.160
TalkEmpahty → M_Bon2	-.006	.098	.901	-.010	.097	.849	-.009	.099	.872
TalkEmpahty → M_Expect2	-.205	.145	.005	-.229	.150	.002	-.220	.153	.005
MusicEmpahty → M_Rel2	-.066	.143	.352	-.056	.143	.434	-.071	.146	.343
MusicEmpahty → M_Bon2	-.007	.094	.893	-.001	.093	.984	.006	.096	.918
MusicEmpahty → M_ Expect2	-.013	.144	.855	-.006	.150	.939	-.020	.152	.803
Squared Multiple Correlation (R2):									
M_Epthy2	.028			.033			.050		
M_Rel2	.337			.339			.283		
M_ Expect2	.370			.395			.353		
M_Bon2	.690			.695			.672		
Model Fit Statistics:									
	χ^2/df = .064, *p* = .938, *GFI* = 1.000, *TLI* = 1.056, *CFI* = 1.000, *RMR* = .004, *RMSEA* = .000			χ^2/df = .066, *p* = .936, *GFI* = 1.000, *TLI* = 1.057, *CFI* = 1.000, *RMR* = .005, *RMSEA* = .000			χ^2/df = .010, *p* = .990, *GFI* = 1.000, *TLI* = 1.070, *CFI* = 1.000, *RMR* = .002, *RMSEA* = .000		

Table 35

Path Analysis of Structural Model with Session 3 data

	N3 = 121			*N3 without MOutlier* = 117			*n3* = 112		
	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*
Hypothesized Relationships									
H1a: *TalkEmpahty → M_Epthy3*	.253	.161	.004	.249	.159	.006	.282	.159	.002
H1b: *MusicEmpahty → M_Epthy3*	-.022	.161	.805	-.033	.159	.711	-.037	.159	.680
H2a: *M_Epthy3 → M_Bon3*	.467	.075	.000	.429	.079	.000	.451	.076	.000
H3a: *M_Epthy3→ M_Rel3*	.649	.093	.000	.636	.098	.000	.626	.100	.000
H3b: *M_Epthy3 → M_Expect3*	.420	.088	.000	.470	.087	.000	.438	.088	.000
H4a: *M_Rel3 → M_Bon3*	.507	.054	.000	.518	.054	.000	.491	.053	.000
H4b: *M_Expect3 → M_Bon3*	-.028	.057	.000	.003	.061	.002	.000	.050	.996
Other Relationships									
TalkEmpahty → M_Rel3	-.039	.169	.586	-.042	.172	.574	.004	.175	.962
TalkEmpahty → M_Bon3	-.142	.104	.005	-.133	.107	.015	-.093	.104	.109
TalkEmpahty → M_Expect3	-.266	.160	.002	-.316	.153	.000	-.326	.155	.005

MusicEmpahty → M_Rel3	.022	.163	.754	.031	.167	.673	.029	.168	.695
MusicEmpahty → M_Bon3	-.050	.097	.294	-.062	.099	.219	-.080	.095	.131
MusicEmpahty → M_ Expect3	.094	.155	.249	.105	.148	.192	.124	148	.139
Squared Multiple Correlation (R2):									
M_Epthy3	.065			.063			.081		
M_Rel3	.198			.254			.229		
M_ Expect3	.410			.392			.393		
M_Bon3	.732			.710			.698		
Model Fit Statistics:									
	χ^2/df = .041, *p* = .959, *GFI* = 1.000, *TLI* = 1.060, *CFI* = 1.000, *RMR* = .002, *RMSEA* = .000			χ^2/df = .150, *p* = .861, *GFI* = .999, *TLI* = 1.056, *CFI* = 1.000, *RMR* = .006, *RMSEA* = .000			χ^2/df = .326, *p* = .722, *GFI* = .998, *TLI* = 1.048, *CFI* = 1.000, *RMR* = .011, *RMSEA* = .000		

Mediating effects of Structural Model

In the last sub-section, we focused on how one construct could directly influence another construct in a SEM model. In this sub-section, we investigated whether and how the third (or more) variable intervened on the influence of the two constructs (Hair et al. 2009). There were serial mediators and parallel mediators in our structural model (Figure 7). To test the indirect effect through the specific path, we denoted all the relationships we were concerned within AMOS before running the SEM analysis. All the paths shown on the left sides on Table 36 – 38 and Mediation analysis summaries are presented on the right side of these tables. The number in the brackets represented the significant level of the direct effect of the independent variable on the dependent variable.

With *N1* = 139 (upper side of Table 36), we first checked the mediation role of *M_Epthy1* on *M_Rel1* and *M_Expect1* respectively. The results revealed insignificant indirect effects of *MusicEmpathy* on *M_Rel1* (*b*= .048, *p* = .564) and on *M_Expect1* (*b*= .044, *p* = .575) through *M_Epthy1*, as well as *TalkEmpathy* on *M_Rel1* (*b*= .019, *p* = .888) and on *M_Expect1* (*b*= .010, *p* = .884) through *M_Epthy1*. We also assessed the serial mediating role of 1) *M_Epthy1* and *M_Rel1 and* 2) *M_Epthy1* and *M_Expect1* on the relationships between empathy forms (i.e., *MusicEmpathy* and *TalkEmpathy*) and *M_Bon1*. However, no significant mediated effect was found. Hence, no variable was found to mediate the relationships between the empathy forms (*MusicEmpathy* and *TalkEmpathy*) and the experiential variables (*M_Bon1*, *M_Expect1*, *M_Rel1*). We were also interested in how the perceived empathy (*M_Epthy1*) affect the *M_Bon1* through *M_Rel1* or *M_Expect1*. The results indicated a positive and significant mediating impact of *M_Rel1* on the relationship between *M_Epthy1* and *M_Bon1* (*b*= .234, *p* = .000). However, the mediating role of *M_Expect1* was insignificant (*b*= .061, *p* = .108). Furthermore, the direct effect of *M_Epthy1* on *M_Bon1* in presence of the mediator was also found significant (*b* = .258, *p* = .001). Hence, *M_Rel1* partially mediated the relationship between *M_Epthy1* and *M_Bon1*. With *N1 without MOutlier* = 130 and *n1* = 125, though the coefficients changed a bit, all the relationships remained the same significance level as found in the mediation analysis results with *N1* = 139. The mediation analysis summary with Session 1 data is presented in Table 36.

With Session 2 data, the partial mediating role of *M_Rel1* on the path from *M_Epthy1* to *M_Bon1* was found (*N1* = 139: *b* = .355, *p* = .000; *N1 without MOutliers* = 130: *b* = .317, *p* = .000; n1 = 125: *b* = .298, *p* = .000). The significant direct effect of *TalkEmpthy* on *M_Epthy1* with Session 2 data wasalso found, consistent with the finding of the path analysis with Session 2 data (Table 34). The mediation analysis summary with Session 2 data is presented in Table 37.

More significant mediators emerged when we ran the SEM analysis with Session 3 data. With all observations (*N3* = 121) and without those who felt extremely negative and lonely (*N3 without MOutliers* = 117), we found *M_Epthy3* fully mediated the relationship between *TalkEmpahty* and *M_Rel3*, and it also partially mediated the relationship between *TalkEmpahty* and *M_Expect3*. Interesting, in the path *TalkEmpahty →M_Epthy3→M_ Expect3*, the direct effect of *TalkEmpahty* on *M_Expect3* had a positive influence, but the indirect effect had a negative influence on *M_Expect3*. It indicated that the participants might have less negative valence on the talking robot when they recognized Zora were trying to show empathy to them. It could be the answer to the question we left in the last sub-section (i.e., Could perceiving Zora tried to do goodness compensate the disappointment Zora talking induced?). Besides, with *N3*

= 121 and *N3 without MOutliers* = 117, the results also revealed a significant and positive mediating role of the serial of *M_Epthy3* and *M_Rel3* on the relationship between *TalkEmpathy* and *M_Bon3*. It was a partial and competitive mediation as *Talkempathy* had a negative and significant impact on *M_Bon3*. Consistent with the mediation effect found with the Session 1 and 2 data, the partial mediating role of *M_Rel3* on the path from *M_Epthy3* to *M_Bon3* remained significant. When we ran the mediation analysis with *n3* = 112, a serial of *M_Epthy3* and *M_Expect3* played a full mediated effect between *TalkEmpathy* and *M_Bon3*, because the direct effect of *TalkEmapthy* and *M_Bon3* went away. The complex mediating effect in Session 3 data came from the negative effect of *TalkEmpathy* on *M_Bon3* and *M_Expect3*, as well as the negative effect of *M_Expect3* on *M_Bon3* (Check Table 35). The mediation analysis summary with Session 3 data is presented in Table 38.

Table 36

Mediating effects of structural model with Session 1 data

Relationships	Direct Effect	Indirect Effect	Confidence Interval Low	Confidence Interval High	ρ-value	Conclusion
N1 = 139						
MusicEmpahty → *M_Epthy1*→ *M_Rel1* (H1a*H3a)	.082	.048	-.117	.232	.564	No mediation
MusicEmpahty → *M_Epthy1*→ *M_ Expect1* (H1a*H3b)	.018	.044	-.112	.203	.575	No mediation
MusicEmpahty → *M_Epthy1*→ *M_Rel1*→ *M_Bon1* (H1a*H3a*H4a)	.008	.019	-.046	.096	.546	No mediation
MusicEmpahty → *M_Epthy1*→ *M_ Expect1*→*M_Bon1* (H1a*H3b*H4b)		.005	-.010	.043	.366	
TalkEmpathy → *M_Epthy1*→ *M_Rel1* (H1b*H3a)	.019	.011	-.154	.198	.888	No mediation
TalkEmpathy →*M_Epthy1*→*M_ Expect1* (H1b*H3b)	-.080	.010	-.142	.183	.884	No mediation
TalkEmpahty → *M_Epthy1*→ *M_Rel1*→ *M_Bon1* (H1b*H3a*H4a)	-.010	.005	-.063	.080	.882	No mediation
TalkEmpahty → *M_Epthy1*→ *M_ Expect1*→ *M_Bon1* (H1b*H3a*H4a)		.001	-.017	.032	.736	
M_Epthy1→ *M_Rel1*→ *M_Bon1* (H3a*H4a)	.258 (.001)	.234	.157	.343	.000	Partial mediation
M_Epthy1→ *M_ Expect1*→*M_Bon1* (H3b*H4b)		.061	-.015	.158	.108	No mediation
N1 without MOutlier = 130						
MusicEmpahty → *M_Epthy1*→ *M_Rel1* (H1a*H3a)	-.229	.032	-.143	.233	.721	No mediation

MusicEmpahty → M_Epthy1→ M_ Expect1 (H1a*H3b)	-.033	.029	-.139	.201	.719	No mediation
MusicEmpahty → M_Epthy1→ M_Rel1→ M_Bon1 (H1a*H3a*H4a)	.010	.013	-.059	.099	.700	No mediation
MusicEmpahty → M_Epthy1→ M_ Expect1→M_Bon1 (H1a*H3b*H4b)		.003	-.011	.038	.459	
TalkEmpahty → M_Epthy1→ M_Rel1 (H1b*H3a)	.003	.005	-.166	.197	.940	No mediation
TalkEmpahty →M_Epthy1→M_ Expect1 (H1b*H3b)	-.125	.004	-.153	.181	.934	No mediation
TalkEmpahty → M_Epthy1→ M_Rel1→ M_Bon1 (H1b*H3a*H4a)	.024	.002	-.070	.080	.937	No mediation
TalkEmpahty → M_Epthy1→ M_ Expect1→ M_Bon1 (H1b*H3a*H4a)		.000	-.018	.027	.837	
M_Epthy1→ M_Rel1→ M_Bon1 (H3a*H4a)	.252 (.004)	.252	.164	.372	.000	Partial mediation
M_Epthy1→ M_ Expect1→M_Bon1 (H3b*H4b)		.051	-.035	.146	.224	No mediation
	n1 = 125					
MusicEmpahty → M_Epthy1→ M_Rel1 (H1a*H3a)	-.249	-.015	-.174	.142	.807	No mediation
MusicEmpahty → M_Epthy1→ M_ Expect1 (H1a*H3b)	-.041	-.014	-.173	.131	.802	No mediation
MusicEmpahty → M_Epthy1→ M_Rel1→ M_Bon1 (H1a*H3a*H4a)	.016	-.006	-.071	.057	.795	No mediation
MusicEmpahty → M_Epthy1→ M_ Expect1→M_Bon1 (H1a*H3b*H4b)		-.002	-.032	.013	.616	
TalkEmpahty → M_Epthy1→ M_Rel1 (H1b*H3a)	-.090	-.057	-.229	.091	.446	No mediation
TalkEmpahty →M_Epthy1→M_ Expect1 (H1b*H3b)	-.174	-.053	-.215	.084	.435	No mediation
TalkEmpahty → M_Epthy1→ M_Rel1→ M_Bon1 (H1b*H3a*H4a)	.000	-.023	-.096	.033	.419	No mediation
TalkEmpahty → M_Epthy1→ M_ Expect1→ M_Bon1 (H1b*H3a*H4a)		-.006	-.048	.006	.270	
M_Epthy1→ M_Rel1→ M_Bon1 (H3a*H4a)	.247 (.004)	.213	.135	.322	.000	Partial mediation
M_Epthy1→ M_ Expect1→M_Bon1 (H3b*H4b)		.054	-.024	.152	.160	No mediation

Note: Unstandardized coefficients reported. Values in parentheses are t-values. Bootstrap sample = 5,000 with replacement.
The number in the brackets represented the significant level of the direct effect of the independent variable on the dependent variable.

Table 37

Mediating effects of structural model with Session 2 data

Relationships	Direct Effect	Indirect Effect	Confidence Interval Low	Confidence Interval High	ρ-value	Conclusion
	N2 = 129					
MusicEmpahty → M_Epthy2→ M_Rel2 (H1a*H3a)	-.133	-.113	-.308	.095	.272	No mediation
MusicEmpahty → M_Epthy2→ M_ Expect2 (H1a*H3b)	-.026	-.109	-.322	.081	.252	No mediation

MusicEmpahty → M_Epthy2→ M_Rel2→ M_Bon2 (H1a*H3a*H4a)	-.013	-.060	-.174	.046	.253	No mediation
MusicEmpahty → M_Epthy2→ M_ Expect2→M_Bon2 (H1a*H3b*H4b)		-.008	-.053	.005	.210	
TalkEmpahty → M_Epthy2→ M_Rel2 (H1b*H3a)	.174	.162	-.036	.372	.106	No mediation
TalkEmpahty →M_Epthy2→M_ Expect2 (H1b*H3b)	-.408 (.004)	.158	-.031	.372	.101	No mediation
TalkEmpahty → M_Epthy2→ M_Rel2→ M_Bon2 (H1b*H3a*H4a)	-.012	.086	-.017	.210	.098	No mediation
TalkEmpahty → M_Epthy2→ M_ Expect2→ M_Bon2 (H1b*H3a*H4a)		.012	-.005	.060	.178	
M_Epthy2→ M_Rel2→ M_Bon2 (H3a*H4a)	.334 (.001)	.355	.252	.487	.000	Partial mediation
M_Epthy2→ M_ Expect2→M_Bon2 (H3b*H4b)		.049	-.040	.138	.266	No mediation
N2 without MOutlier = 122						
MusicEmpahty → M_Epthy2→ M_Rel2 (H1a*H3a)	-.112	-.137	-.349	.075	.191	No mediation
MusicEmpahty → M_Epthy2→ M_ Expect2 (H1a*H3b)	-.011	-.130	-.363	.066	.181	No mediation
MusicEmpahty → M_Epthy2→ M_Rel2→ M_Bon2 (H1a*H3a*H4a)	-.002	-.064	-.184	.030	.169	No mediation
MusicEmpahty → M_Epthy2→ M_ Expect2→M_Bon2 (H1a*H3b*H4b)		-.007	-.052	.006	.247	
TalkEmpahty → M_Epthy2→ M_Rel2 (H1b*H3a)	.196	.168	-.045	.394	.112	No mediation
TalkEmpahty →M_Epthy2→M_ Expect2 (H1b*H3b)	-.460 (.004)	.159	-.038	.385	.108	No mediation
TalkEmpahty → M_Epthy2→ M_Rel2→ M_Bon2 (H1b*H3a*H4a)	-.018	.079	-.018	.194	.103	No mediation
TalkEmpahty → M_Epthy2→ M_ Expect2→ M_Bon2 (H1b*H3a*H4a)		.009	-.008	.057	.252	
M_Epthy2→ M_Rel2→ M_Bon2 (H3a*H4a)	.391 (.001)	.317	.223	.439	.000	Partial mediation
M_Epthy2→ M_ Expect2→M_Bon2 (H3b*H4b)		.036	-.050	.124	.396	No mediation
n2 = 119						
MusicEmpahty → M_Epthy2→ M_Rel2 (H1a*H3a)	-.138	-.157	-.345	.036	.113	No mediation
MusicEmpahty → M_Epthy2→ M_ Expect2 (H1a*H3b)	-.038	-.147	-.365	.029	.103	No mediation
MusicEmpahty → M_Epthy2→ M_Rel2→ M_Bon2 (H1a*H3a*H4a)	.010	-.074	-.183	.013	.101	No mediation
MusicEmpahty → M_Epthy2→ M_ Expect2→M_Bon2 (H1a*H3b*H4b)		-.009	-.055	.007	.246	
TalkEmpahty → M_Epthy2→ M_Rel2 (H1b*H3a)	.205	.180	-.010	.372	.067	No mediation
TalkEmpahty →M_Epthy2→M_ Expect2 (H1b*H3b)	-.425 (.007)	.168	-.006	.382	.059	No mediation
TalkEmpahty → M_Epthy2→ M_Rel2→ M_Bon2 (H1b*H3a*H4a)	-.016	.085	-.002	.191	.055	No mediation
TalkEmpahty → M_Epthy2→ M_ Expect2→ M_Bon2 (H1b*H3a*H4a)		.010	-.009	.053	.245	
M_Epthy2→ M_Rel2→ M_Bon2 (H3a*H4a)	.395 (.000)	.298	.205	.416	.000	Partial mediation
M_Epthy2→ M_ Expect2→M_Bon2 (H3b*H4b)		.035	-.047	.121	.365	No mediation

Note: Unstandardized coefficients reported. Values in parentheses are t-values. Bootstrap sample = 5,000 with replacement.

The number in the brackets represented the significant level of the direct effect of the independent variable on the dependent variable.

Table 37

Mediating effects of structural model with Session 3 data

Relationships	Direct Effect	Indirect Effect	Confidence Interval Low	Confidence Interval High	ρ-value	Conclusion
N3 = 121						
MusicEmpahty → M_Epthy3→ M_Rel3 (H1a*H3a)	.051	-.033	-.290	.232	.787	No mediation
MusicEmpahty → M_Epthy3→ M_ Expect3 (H1a*H3b)	.179	-.017	-.173	.118	.762	No mediation
MusicEmpahty → M_Epthy3→ M_Rel3→ M_Bon3 (H1a*H3a*H4a)	-.102	-.015	-.128	.107	.784	No mediation
MusicEmpahty → M_Epthy3→ M_ Expect3→M_Bon3 (H1a*H3b*H4b)		.001	-.006	.015	.607	
TalkEmpahty → M_Epthy3→ M_Rel3 (H1b*H3a)	-.092	.383	.112	.682	.005	Full mediation
TalkEmpahty →M_Epthy3→M_ Expect3 (H1b*H3b)	-.504 (.005)	.201	.054	.424	.004	Partial mediation
TalkEmpahty → M_Epthy3→ M_Rel3→ M_Bon3 (H1b*H3a*H4a)	-.290 (.012)	.170	.056	.324	.003	Partial mediation
TalkEmpahty → M_Epthy3→ M_ Expect3→ M_Bon3 (H1b*H3a*H4a)		-.006	-.038	.015	.412	No mediation
M_Epthy3→ M_Rel3→ M_Bon3 (H3a*H4a)	.523 (.001)	.368	.262	.520	.000	Partial mediation
M_Epthy3→ M_ Expect3→M_Bon3 (H3b*H4b)		-.013	-.062	.041	.531	No mediation
N3 without MOutlier = 117						
MusicEmpahty → M_Epthy3→ M_Rel3 (H1a*H3a)	.070	-.049	-.303	.219	.710	No mediation
MusicEmpahty → M_Epthy3→ M_ Expect3 (H1a*H3b)	.193	-.029	-.197	.126	.681	No mediation
MusicEmpahty → M_Epthy3→ M_Rel3→ M_Bon3 (H1a*H3a*H4a)	-.121	-.021	-.136	.102	.687	No mediation
MusicEmpahty → M_Epthy3→ M_ Expect3→M_Bon3 (H1a*H3b*H4b)		.000	-.013	.009	.882	
TalkEmpahty → M_Epthy3→ M_Rel3 (H1b*H3a)	-.097	.364	.104	.669	.010	Full mediation
TalkEmpahty →M_Epthy3→M_ Expect3 (H1b*H3b)	-.683 (.001)	.216	.063	.437	.007	Partial mediation
TalkEmpahty → M_Epthy3→ M_Rel3→ M_Bon3 (H1b*H3a*H4a)	-.261 (.024)	.160	.048	.312	.007	Partial mediation
TalkEmpahty → M_Epthy3→ M_ Expect3→ M_Bon3 (H1b*H3a*H4a)		.001	-.027	.032	.914	No mediation
M_Epthy3→ M_Rel3→ M_Bon3 (H3a*H4a)	.475 (.000)	.365	.251	.516	.000	Partial mediation
M_Epthy3→ M_ Expect3→M_Bon3 (H3b*H4b)		.001	-.054	.067	.950	No mediation
n3 = 112						

MusicEmpahty → M_Epthy3→ M_Rel3 (H1a*H3a)	.066	-.053	-.299	.219	.686	No mediation
MusicEmpahty → M_Epthy3→ M_ Expect3 (H1a*H3b)	.220	-.029	-.184	.116	.666	No mediation
MusicEmpahty → M_Epthy3→ M_Rel3→ M_Bon3 (H1a*H3a*H4a)	-.144	-.021	-.120	.087	.664	No mediation
MusicEmpahty → M_Epthy3→ M_ Expect3→M_Bon3 (H1a*H3b*H4b)		.000	-.010	.010	.963	
TalkEmpahty → M_Epthy3→ M_Rel3 (H1b*H3a)	.008	.401	.132	.708	.003	Full mediation
TalkEmpahty →M_Epthy3→M_ Expect3 (H1b*H3b)	-.581 (.001)	.220	.070	.444	.002	Partial mediation
TalkEmpahty → M_Epthy3→ M_Rel3→ M_Bon3 (H1b*H3a*H4a)	-.167	.157	.054	.303	.002	Full mediation
TalkEmpahty → M_Epthy3→ M_ Expect3→ M_Bon3 (H1b*H3a*H4a)		.000	-.029	.027	.998	No mediation
M_Epthy3→ M_Rel3→ M_Bon3 (H3a*H4a)	.465 (.000)	.317	.213	.450	.000	Partial mediation
M_Epthy3→ M_ Expect3→M_Bon3 (H3b*H4b)		.000	-.049	.055	.997	No mediation

Note: Unstandardized coefficients reported. Values in parentheses are t-values. Bootstrap sample = 5,000 with replacement.
The number in the brackets represented the significant level of the direct effect of the independent variable on the dependent variable.

Moderation effects of Structural Model

In our complete hypothesized theory model, we assumed that the level of loneliness (*M_Lon#*) could alter the direct influence of the perceived empathy (*M_Epthy#*) on relevance (*M_Rel#*) and valence (*M_Expect#*). This sub-section assessed with 2 sessions data the moderating role of *M_Lon#* on the relationships between 1) *M_Epthy#* and *M_Rel#* and 2) *M_Epthy#* and *M_Expect#*. The results revealed a negative and significant moderating impact of *M_Lon1* on the relationship between *M_Rel1* and *M_Epthy1* (b = -.921, *s.e.* = .104, p = .031). However, the moderated effect of *M_Lon#* was insignificant with other datasets. It might be because in the datasets without MOutlier (e.g., *N1 without MOutlier* = 130) and without both MOutliesr and EOutliers (e.g., *n1* = 125), the lonely people were excluded and thus the moderated effect was gone away. Regarding the datasets with all observations (*N2* = 129, *N3*=121), we had no evidence to explain why the moderated effects disappeared. What interested us most was that the direct effect of *M_Epthy#* on *M_Expect#* became not significant when we added the moderator *M_Lon#* in the model. We had checked there was no multicollinearity problem with *M_Lon#* and *M_Epthy#*.

Table 38

Moderation effects with Session 1 data

	N1 = 139			*N1 without MOutlier* = 130			*n1* = 125		
Relationships	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*
M_Lon1 → M_Rel1	.809	.369	.013	.440	.403	.167	.570	.407	.094
M_Lon1× M_Epthy1 → M_Rel1	-.921	.104	.031	-.521	.116	.226	-.654	.117	.144
M_Epthy1→ M_Rel1	1.069	.319	.000	.875	.346	.003	.893	.347	.003
(-1SD) M_Lon1 on M_Rel1				-	-	-	-	-	-
(SD) M_Lon1 on M_Rel1				-	-	-	-	-	-
(+SD) M_Lon1 on M_Rel1				-	-	-	-	-	-
M_Lon1 → M_Expect1	.237	.369	.480	-.091	.387	.778	-.194	.395	.575
M_Lon1× M_Epthy1 → M_Expect1	-.453	.102	.300	.017	.111	.968	.141	.113	.757
M_Epthy1→ M_Expect1	.766	.311	.007	.496	.332	.103	.367	.337	.227
(-1SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(+SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-

Table 39

Moderation effects with Session 2 data

	N2 = 129			*N2 without MOutlier* = 122			*n2* = 119		
Relationships	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*
M_Lon1 → M_Rel1	-.001	.077	.998	.084	.078	.726	.071	.284	.781
M_Lon1× M_Epthy1 → M_Rel1	.106	.248	.693	-.027	.252	.923	-.011	.083	.970
M_Epthy1→ M_Rel1	.544	.262	.012	.655	.271	.003	.605	.263	.009
(-1SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(+SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
M_Lon1 → M_Expect1	-.348	.270	.165	-.273	.289	.286	-.156	.301	.571
M_Lon1× M_Epthy1 → M_Expect1	.277	.079	.319	.202	.084	.493	.037	.089	.906
M_Epthy1→ M_Expect1	.301	.255	.182	.355	.269	.136	.425	.279	.083
(-1SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(+SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-

Table 40

Moderation effects with Session 3 data

	N3 = 121			*N3 without MOutlier* = 117			*n3* = 112		
Relationships	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*	Std. Estimates	*S.E.*	*p*
M_Lon1 → M_Rel1	.293	.270	.197	.455	.374	.103	.399	.381	164
M_Lon1× M_Epthy1 → M_Rel1	-.341	.079	.226	-.591	.101	.099	-.491	.104	.171
M_Epthy1→ M_Rel1	.868	.255	.000	.986	.301	.000	.935	.307	.000
(-1SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
(+SD) M_Lon1 on M_Rel1	-	-	-	-	-	-	-	-	-
M_Lon1 → M_Expect1	.387	.265	.157	.157	.353	.633	.325	.351	.334

M_Lon1× M_Epthy1 → M_Expect1	-.605	.077	.075	-.376	.095	.373	-.679	.096	.106
M_Epthy1→ M_Expect1	.746	.250	.002	.622	.284	.022	.762	.283	.006
(-1SD) *M_Lon1* on *M_Rel1*	-	-	-	-	-	-	-	-	-
(SD) *M_Lon1* on *M_Rel1*	-	-	-	-	-	-	-	-	-
(+SD) *M_Lon1* on *M_Rel1*	-	-	-	-	-	-	-	-	-

Acknowledgements

This research was supported by the project Negative-mood reduction among HK youth with robot PAL (Personal Avatar for Life) of the Artificial Intelligence in Design Laboratory in Hong Kong (grant number: AiDLab RP2P3).

References

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Broersen, T.M.A.(2018). De invloed van affectieve binding met een servicerobot op debeleving van gastvrijheid in een hotelomgeving (masterthesis). *Geraadpleegd*. http://www.ubvu.vu.nl/pub/index_oclc.cfm?SearchObjectId=8&objectid=109&fctN=st50&fctIncV=1&fctExcV=+#Y.

Cattell, R. B. (1966). The screen test for the number of factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cohen, J. (1988). *Statistical Power Analysis for the behavioral sciences*. L. Erlbaum Associates.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/bf02310555

Cyranowski, J. M., Zill, N., Bode, R., Butt, Z., Kelly, M. A., Pilkonis, P. A., Salsman, J. M., & Cella, D. (2013). Assessing social support, companionship, and distress: National Institute of Health (NIH) toolbox adult social relationship scales. *Health Psychology, 32*(3), 293–301. https://doi.org/10.1037/a0028586

De Jong Gierveld, J., & Van Tilburg, T. (2010). The de jong gierveld short scales for emotional and social loneliness: Tested on data from 7 countries in the UN Generations and gender surveys. *European Journal of Ageing*, *7*(2), 121–130. https://doi.org/10.1007/s10433-010-0144-6

Field, A. (2014). *Discovering statistics using Ibm Spss statistics*. SAGE Publications.

Hair, J. F. (2009). *Multivariate Data Analysis: A global perspective*. Prentice Hall.

Haile, L., Gallagher, M., & J. Robertson, R. (2014). Perceived exertion laboratory manual: From standard practice to contemporary application. *Perceived Exertion Laboratory Manual*, 11–20. https://doi.org/10.1007/978-1-4939-1917-8_2

Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2019). *Multivariate Data Analysis*. Cengage.

Hoorn, J. F., & Konijn, E. A. (2003). Perceiving and experiencing fictional characters: An integrative account1. *Japanese Psychological Research*, *45*(4), 250–268. https://doi.org/10.1111/1468-5884.00225

Hoyle, R. (2011). Structural equation modeling for social and personality psychology. https://doi.org/10.4135/9781446287965

Hu, L.-tze, & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. https://doi.org/10.1037/1082-989x.3.4.424

Kaiser, H. F. (1960). The application of electronic computers to Factor Analysis. *Educational and Psychological Measurement*, *20*(1), 141–151. https://doi.org/10.1177/001316446002000116

Kaiser, H. F. (1974). An index of Factorial Simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/bf02291575

Kline, R. B. (2016). *Principles and practice of structural equation modeling*. The Guilford Press.

Konijn, E. A., & Hoorn, J. F. (2005). Some like it bad: Testing a model for perceiving and experiencing fictional characters. *Media Psychology*, *7*(2), 107–144. https://doi.org/10.1207/s1532785xmep0702_1

Konijn, E.A. & Hoorn, J.F., 2017. ‘Parasocial Interaction and Beyond: Media Personae and Affective Bonding’. *The International Encyclopedia of Media Effects*, pp.1–15.

Konijn, E. A., Jansen, B., Mondaca Bustos, V., Hobbelink, V. L., & Preciado Vanegas, D. (2021). Social Robots for (second) language learning in (migrant) primary school children. *International Journal of Social Robotics*, *14*(3), 827–843. https://doi.org/10.1007/s12369-021-00824-3

Nunnally, J. C. (1975). Psychometric theory. 25 years ago and now. *Educational Researcher*, *4*(10), 7. https://doi.org/10.2307/1175619

Paauwe, R. A., Hoorn, J. F., Konijn, E. A., & Keyson, D. V. (2015). Designing robot embodiments for social interaction: Affordances topple realism and aesthetics. *International Journal of Social Robotics*, *7*(5), 697–708. https://doi.org/10.1007/s12369-015-0301-3

Salkind, N. J. (Ed.). (2010). Multivariate analysis of variance (MANOVA). *Encyclopedia of Research Design*. https://doi.org/10.4135/9781412961288.n257

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting, recruiting: A brief introduction to running behavioral experiments online. https://doi.org/10.31234/osf.io/tr76d

Smith, M. D., & Hoyle, R. H. (1996). Structural equation modeling: Concepts, issues, and applications. *The Statistician*, *45*(2), 267. https://doi.org/10.2307/2988418

Soper, D. (2022). *Calculator: A-priori sample size for structural equation models*. Free A-priori Sample Size Calculator for Structural Equation Models - Free Statistics Calculators. Retrieved June 24, 2022, from https://www.danielsoper.com/statcalc/calculator.aspx?id=89

Tabachnick, B. G., & Fidell, L. S. (2001). *Computer-Assisted Research Design and analysis*. Allyn and Bacon.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/bf02291170

Ullman, J. B. (2006). Structural equation modeling: Reviewing the basics and moving forward. *Journal of Personality Assessment*, *87*(1), 35–50. https://doi.org/10.1207/s15327752jpa8701_03

van Vugt, H.C., Hoorn, J.F. & Konijn, E.A., 2009. ‘Interactive engagement with embodied agents: an empirically validated framework’. *Computer Animation and Virtual Worlds*, 20(2-3), pp.195–204.

Waytz, A., & Epley, N. (2012). Social Connection enables dehumanization. *Journal of Experimental Social Psychology*, *48*(1), 70–76. https://doi.org/10.1016/j.jesp.2011.07.012

Westland, J. C. (2010). Erratum to “lower bounds on sample size in structural equation modeling” [electron. Commerce Res. appl. 9 (6) (2010) 476–487]. *Electronic Commerce Research and Applications*, *9*(6), 476–487. https://doi.org/10.1016/j.elerap.2012.06.001

Appendix 1:

Table 1

Correlation matrix for experiential variables and covariates (N1 without MOutliers =130)

Scale	*M_Bon1*	*M_PerEmp1*	*M_Rel1*	*M_Expect1*	*M_DesReal1*
			Condition 1		
M_Bon1	1				
M_PerEmp1	.574**	1			
M_Rel1	.731**	.617**	1		
M_Expect1	.686**	.663**	.595**	1	
M_DesReal1	.639**	.543**	.815**	.627**	1
			Condition 2		
M_Bon1	1				
M_PerEmp1	.389*	1			
M_Rel1	.592**	.530**	1		
M_Expect1	.329	.339	.403*	1	
M_DesReal1	.607**	.526**	.711**	.191	1
			Condition 3		
M_Bon1	1				
M_PerEmp1	.602**	1			
M_Rel1	.776**	.594**	1		
M_Expect1	.207	.443*	.296	1	
M_DesReal1	.641**	.574**	.755**	.184	1
			Condition 4		
M_Bon1	1				
M_PerEmp1	.513**	1			
M_Rel1	.439**	.372*	1		
M_Expect1	.417*	.632**	.100	1	
M_DesReal1	.462**	.490**	.538**	.496**	1

Note. Pearson correlation coefficients based on 5000 bootstrap samples are presented above (*N* =139).
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.000 level (2-tailed).

Table 2

Box's Test of Equality of Covariance Matrices[a] (N1 without MOutliers =130)

Box's M	62.602
F	1.291
df1	45
df2	37808.829
Sig.	.091

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3

Levene's Test of Equality of Error Variances[a] (N1 without MOutliers =130)

	F	df1	df2	Sig.
M_Bon1	2.076	3	126	.107
M_PerEmp1	1.240	3	126	.298

M_Rel1	.548	3	126	.650
M_Expect1	.235	3	126	.872
M_DesReal1	.767	3	126	.514

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on *MusicEmpathy* and *TalkEmpathy* (*N1 without MOutliers =130*)

Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	.714	9.553	5.000	119.000	.000	.286
M_Lon1	.922	2.011	5.000	119.000	.082	.078
M_EmoP1	.873	3.455	5.000	119.000	.006	.127
M_EmoN1	.913	2.281	5.000	119.000	.051	.087
MusicEmpathy	.976	.585	5.000	119.000	.712	.024
TalkEmpathy	.992	.198	5.000	119.000	.963	.008
MusicEmpathy TalkEmpathy*	.939	1.554	5.000	119.000	.179	.061

Table 5 Tests of Between-Subjects Effects on experiential variables (N1 =139, excluded mental state outliers)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon1	*M_Bon1*	1	.270	.369	.545	.003
	M_PerEmp1	1	.541	.756	.386	.006
	M_Rel1	1	1.829	1.799	.182	.014
	M_Expect1	1	1.826	1.964	.164	.016
	M_DesReal1	1	.112	.152	.697	.001
M_EmoP1	*M_Bon1*	1	6.038	8.246	.005	.063
	M_PerEmp1	1	.761	1.063	.305	.009
	M_Rel1	1	2.445	2.406	.123	.019
	M_Expect1	1	1.983	2.133	.147	.017
	M_DesReal1	1	1.364	1.856	.176	.015
M_EmoN1	*M_Bon1*	1	.066	.090	.765	.001
	M_PerEmp1	1	3.345	4.676	.033	.037
	M_Rel1	1	.722	.711	.401	.006
	M_Expect1	1	.799	.859	.356	.007
	M_DesReal1	1	.443	.603	.439	.005
MusicEmpathy	*M_Bon1*	1	.002	.002	.963	.000
	M_PerEmp1	1	.160	.224	.637	.002
	M_Rel1	1	1.200	1.181	.279	.010
	M_Expect1	1	.004	.004	.948	.000
	M_DesReal1	1	.189	.257	.613	.002
TalkEmpathy	*M_Bon1*	1	.024	.033	.856	.000
	M_PerEmp1	1	.003	.005	.945	.000
	M_Rel1	1	.014	.014	.905	.000
	M_Expect1	1	.402	.432	.512	.004
	M_DesReal1	1	.083	.112	.738	.001
*MusicEmpathy * TalkEmpathy*	*M_Bon1*	1	4.106	5.607	.019	.044
	M_PerEmp1	1	3.533	4.938	.028	.039
	M_Rel1	1	2.819	2.774	.098	.022
	M_Expect1	1	.390	.420	.518	.003
	M_DesReal1	1	1.842	2.507	.116	.020

Appendix 2:

Table 1 Correlation matrix for experiential variables and covariates (n1 =125)

Scale	*M_Bon1*	*M_PerEmp1*	*M_Rel1*	*M_Expect1*	*M_DesReal1*
		Condition 1			
M_Bon1	1				
M_PerEmp1	.385*	1			
M_Rel1	.619**	.413*	1		
M_Expect1	.578**	.527**	.440*	1	
M_DesReal1	.499**	.303	.728**	.518**	1
		Condition 2			
M_Bon1	1				
M_PerEmp1	.370*	1			
M_Rel1	.506**	.504**	1		
M_Expect1	.308	.210	.352*	1	
M_DesReal1	.495**	.586**	.667**	.190	1
		Condition 3			
M_Bon1	1				
M_PerEmp1	.602**	1			
M_Rel1	.776**	.594**	1		
M_Expect1	.207	.443*	.296	1	
M_DesReal1	.641**	.574**	.755**	.184	1
		Condition 4			
M_Bon1	1				
M_PerEmp1	.389*	1			
M_Rel1	.373*	.282	1		
M_Expect1	.400*	.660**	.066	1	
M_DesReal1	.432*	.469**	.517**	.484**	1

Note. Pearson correlation coefficients based on 5000 bootstrap samples are presented above (*N* =139).
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.000 level (2-tailed).

Table 2 Box's Test of Equality of Covariance Matrices[a] (n1 =125)

Box's M	65.168
F	1.339
df1	45
df2	34249.783
Sig.	.064

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (n1 =125)

	F	df1	df2	Sig.
M_Bon1	3.283	3	121	.023
M_PerEmp1	2.606	3	121	.055
M_Rel1	.440	3	121	.725
M_Expect1	.457	3	121	.713
M_DesReal1	1.083	3	121	.359

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (*N1 without MOutliers* =130)

Effect	**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.690	10.224	5.000	114.000	.000	.310
M_Lon1	.916	2.078	5.000	114.000	.073	.084
M_EmoP1	.878	3.171	5.000	114.000	.010	.122

M_EmoN1	.900	2.531	5.000	114.000	.033	.100
MusicEmpathy	.973	.633	5.000	114.000	.675	.027
TalkEmpathy	.982	.409	5.000	114.000	.842	.018
MusicEmpathy TalkEmpathy*	.922	1.925	5.000	114.000	.096	.078

Table 5 Tests of Between-Subjects Effects on experiential variables (*n1* =125)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon1	*M_Bon1*	1	.448	.676	.412	.006
	M_PerEmp1	1	.644	1.026	.313	.009
	M_Rel1	1	2.034	2.271	.135	.019
	M_Expect1	1	1.865	2.205	.140	.018
	M_DesReal1	1	.057	.090	.765	.001
M_EmoP1	*M_Bon1*	1	5.242	7.917	.006	.063
	M_PerEmp1	1	.634	1.009	.317	.008
	M_Rel1	1	2.190	2.445	.121	.020
	M_Expect1	1	1.798	2.126	.147	.018
	M_DesReal1	1	1.206	1.906	.170	.016
M_EmoN1	*M_Bon1*	1	.075	.114	.737	.001
	M_PerEmp1	1	3.436	5.472	.021	.044
	M_Rel1	1	.803	.897	.346	.008
	M_Expect1	1	.667	.789	.376	.007
	M_DesReal1	1	.490	.774	.381	.007
MusicEmpathy	*M_Bon1*	1	.066	.099	.753	.001
	M_PerEmp1	1	9.011E-5	.000	.990	.000
	M_Rel1	1	2.038	2.275	.134	.019
	M_Expect1	1	.045	.053	.818	.000
	M_DesReal1	1	.491	.776	.380	.007
TalkEmpathy	*M_Bon1*	1	.595	.898	.345	.008
	M_PerEmp1	1	.290	.462	.498	.004
	M_Rel1	1	.513	.573	.451	.005
	M_Expect1	1	1.504	1.779	.185	.015
	M_DesReal1	1	.184	.291	.591	.002
*MusicEmpathy * TalkEmpathy*	*M_Bon1*	1	4.497	6.792	.010	.054
	M_PerEmp1	1	3.508	5.587	.020	.045
	M_Rel1	1	2.499	2.790	.097	.023
	M_Expect1	1	.216	.255	.614	.002
	M_DesReal1	1	1.509	2.385	.125	.020

Table 6 Significant results of Univariate Test (*n1* =125)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2
M_Bon1	*TalkEmpathyN*	*MusicEmpathyN*	*MusicEmpathyY*	.429	.036	4.497	.037
	MusicEpthyN	*TalkEmpathyN*	*TalkEmpathyY*	.523	.011	6.703	.054
M_PerEmp1	*MusicEpthyN*	*TalkEmpathyN*	*TalkEmpathyY*	.436	.029	4.910	.040
M_Rel1	*TalkEmpathyN*	*MusicEmpathyN*	*MusicEmpathyY*	.544	.023	5.342	.043

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 3:

Table 1 Correlation matrix for experiential variables and covariates (*N2* =129)

Scale	M_Bon1	M_PerEmp1	M_Rel1	M_Expect1	M_DesReal1
		Condition 1			
M_Bon2	1				
M_PerEmp2	.524**	1			
M_Rel2	.793**	.513**	1		
M_Expect2	.621**	.735**	.506**	1	
M_DesReal2	.472**	.635**	.729**	.578**	1
		Condition 2			
M_Bon1	1				
M_PerEmp2	.620**	1			
M_Rel2	.805**	.555**	1		
M_Expect2	.469**	.625**	.481**	1	
M_DesReal2	.298	.234	.538**	.064	1
		Condition 3			
M_Bon2	1				
M_PerEmp2	.830**	1			
M_Rel2	.776**	.726**	1		
M_Expect2	.365*	.479**	.291	1	
M_DesReal2	.711**	.570**	.758**	.348	1
		Condition 4			
M_Bon2	1				
M_PerEmp2	.723**	1			
M_Rel2	.760**	.548**	1		
M_Expect2	.473**	.491**	.167	1	
M_DesReal2	.725**	.701**	.536**	.421*	1

Note. Pearson correlation coefficients based on 5000 bootstrap samples are presented above (*N* =129).
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.000 level (2-tailed).

Table 2 Box's Test of Equality of Covariance Matrices[a] (*N2* =127)

Box's M	74.254
F	1.528
df1	45
df2	36561.712
Sig.	.013

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (*N2* =127)

	F	df1	df2	Sig.
M_Bon2	1.289	3	123	.281
M_PerEmp2	.186	3	123	.906
M_Rel2	.078	3	123	.972
M_Expect2	.272	3	123	.845
M_DesReal2	1.336	3	123	.266

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (*N2* =127)

Effect	**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.576	17.065	5.000	116.000	.000	.424
M_Lon2	.931	1.708	5.000	116.000	.138	.069
M_EmoP2	.785	6.362	5.000	116.000	.000	.215
M_EmoN2	.832	4.682	5.000	116.000	.001	.168

MusicEmpathy	.970	.722	5.000	116.000	.609	.030
TalkEmpathy	.893	2.788	5.000	116.000	.020	.107
MusicEmpathy TalkEmpathy*	.931	1.728	5.000	116.000	.134	.069

a. Design: Intercept + M_Lon2 + M_EmoP2 + M_EmoN2 + MusicY + TalkY + MusicY * TalkY

Table 5 Tests of Between-Subjects Effects on experiential variables (*N2* =127)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon1	*M_Bon2*	1	.127	.165	.685	.001
	M_PerEmp2	1	1.879	2.737	.101	.022
	M_Rel2	1	.198	.226	.635	.002
	M_Expect2	1	2.002	2.218	.139	.018
	M_DesReal2	1	.093	.113	.737	.001
M_EmoP1	*M_Bon2*	1	12.878	16.771	.000	.123
	M_PerEmp2	1	2.204	3.210	.076	.026
	M_Rel2	1	16.396	18.720	.000	.135
	M_Expect2	1	.512	.567	.453	.005
	M_DesReal2	1	2.439	2.980	.087	.024
M_EmoN1	*M_Bon2*	1	.004	.005	.945	.000
	M_PerEmp2	1	.091	.132	.717	.001
	M_Rel2	1	5.153	5.883	.017	.047
	M_Expect2	1	3.756	4.161	.044	.034
	M_DesReal2	1	.001	.001	.974	.000
MusicEmpathy	*M_Bon2*	1	.023	.030	.862	.000
	M_PerEmp2	1	.001	.001	.969	.000
	M_Rel2	1	.453	.517	.474	.004
	M_Expect2	1	.195	.216	.643	.002
	M_DesReal2	1	1.757	2.146	.146	.018
TalkEmpathy	*M_Bon2*	1	1.193	1.553	.215	.013
	M_PerEmp2	1	1.104	1.609	.207	.013
	M_Rel2	1	2.342	2.674	.105	.022
	M_Expect2	1	3.320	3.678	.058	.030
	M_DesReal2	1	.162	.198	.657	.002
*MusicEmpathy * TalkEmpathy*	*M_Bon2*	1	3.072	4.001	.048	.032
	M_PerEmp2	1	.297	.432	.512	.004
	M_Rel2	1	1.367	1.561	.214	.013
	M_Expect2	1	.822	.911	.342	.008
	M_DesReal2	1	.519	.634	.427	.005

Table 6 Significant results of Univariate Test (*N2* =127)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2
M_Bon2	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	-511	.025	5.134	.041
M_Rel2	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	-.485	.047	4.044	.033
M_Expect2	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.490	.048	4.005	.032

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 4:

Table 1 Correlation matrix for experiential variables and covariates (*N2 without MOutliers* =122)

Scale	M_Bon1	M_PerEmp1	M_Rel1	M_Expect1	M_DesReal1
		Condition 1			
M_Bon2	1				
M_PerEmp2	.699**	1			
M_Rel2	.779**	.597**	1		
M_Expect2	.764**	.734**	.556**	1	
M_DesReal2	.579**	.637**	.785**	.577**	1
		Condition 2			
M_Bon1	1				
M_PerEmp2	.618**	1			
M_Rel2	.814**	.570**	1		
M_Expect2	.470**	.620**	.519**	1	
M_DesReal2	.299	.236	.539**	.068	1
		Condition 3			
M_Bon2	1				
M_PerEmp2	.844**	1			
M_Rel2	.769**	.728**	1		
M_Expect2	.382*	.505**	.307	1	
M_DesReal2	.736**	.649**	.793**	.360	1
		Condition 4			
M_Bon2	1				
M_PerEmp2	.703**	1			
M_Rel2	.742**	.515**	1		
M_Expect2	.396*	.450**	.101	1	
M_DesReal2	.700**	.676**	.498**	.371*	1

Table 2 Box's Test of Equality of Covariance Matrices[a] (*N2 without MOutliers* =122)

Box's M	73.249
F	1.500
df1	45
df2	32439.450
Sig.	.017

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (*N2 without MOutliers* =120)

	F	df1	df2	Sig.
M_Bon2	2.412	3	116	.070
M_PerEmp2	.066	3	116	.978
M_Rel2	.152	3	116	.928
M_Expect2	.362	3	116	.780
M_DesReal2	1.146	3	116	.334

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (*N2 without MOutliers* =120)

Effect	**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.667	10.906	5.000	109.000	.000	.333
M_Lon2	.955	1.033	5.000	109.000	.402	.045
M_EmoP2	.842	4.084	5.000	109.000	.002	.158
M_EmoN2	.841	4.113	5.000	109.000	.002	.159
MusicEmpathy	.971	.640	5.000	109.000	.669	.029
TalkEmpathy	.885	2.831	5.000	109.000	.019	.115
MusicEmpathy TalkEmpathy*	.928	1.701	5.000	109.000	.140	.072

a. Design: Intercept + M_Lon2 + M_EmoP2 + M_EmoN2 + MusicY + TalkY + MusicY * TalkY

Table 5 Tests of Between-Subjects Effects on experiential variables (*N2 without MOutliers* =120)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon1	*M_Bon2*	1	.130	.177	.675	.002
	M_PerEmp2	1	.567	.815	.369	.007
	M_Rel2	1	.518	.584	.446	.005
	M_Expect2	1	1.469	1.582	.211	.014
	M_DesReal2	1	.004	.004	.947	.000
M_EmoP1	*M_Bon2*	1	6.785	9.199	.003	.075
	M_PerEmp2	1	3.377	4.854	.030	.041
	M_Rel2	1	12.294	13.875	.000	.109
	M_Expect2	1	.270	.291	.591	.003
	M_DesReal2	1	2.388	2.831	.095	.024
M_EmoN1	*M_Bon2*	1	.248	.336	.563	.003
	M_PerEmp2	1	.114	.164	.686	.001
	M_Rel2	1	3.831	4.324	.040	.037
	M_Expect2	1	3.316	3.571	.061	.031
	M_DesReal2	1	.055	.065	.799	.001
MusicEmpathy	*M_Bon2*	1	.024	.033	.856	.000
	M_PerEmp2	1	.000	.001	.979	.000
	M_Rel2	1	.539	.608	.437	.005
	M_Expect2	1	.106	.114	.737	.001
	M_DesReal2	1	1.542	1.828	.179	.016
TalkEmpathy	*M_Bon2*	1	1.065	1.444	.232	.013
	M_PerEmp2	1	1.125	1.618	.206	.014
	M_Rel2	1	1.877	2.118	.148	.018
	M_Expect2	1	3.449	3.715	.056	.032
	M_DesReal2	1	.010	.012	.914	.000
*MusicEmpathy * TalkEmpathy*	*M_Bon2*	1	3.486	4.726	.032	.040
	M_PerEmp2	1	.504	.725	.396	.006
	M_Rel2	1	1.293	1.460	.230	.013
	M_Expect2	1	.406	.437	.510	.004
	M_DesReal2	1	.397	.471	.494	.004

Table 6 Significant results of Univariate Test (*N2 without MOutliers* =120)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2
M_Bon2	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.538	.020	5.543	.047

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 5:

Table 1 Correlation matrix for experiential variables and covariates (*n2* =119)

Scale	*M_Bon1*	*M_PerEmp1*	*M_Rel1*	*M_Expect1*	*M_DesReal1*
			Condition 1		
M_Bon2	1				
M_PerEmp2	.656**	1			
M_Rel2	.745**	.526**	1		
M_Expect2	.735**	.700**	.491**	1	
M_DesReal2	.499**	.559**	.734**	.506**	1
			Condition 2		
M_Bon2	1				
M_PerEmp2	.530**	1			
M_Rel2	.783**	.484**	1		
M_Expect2	.339	.499**	.418*	1	
M_DesReal2	.211	.119	.489**	-.089	1
			Condition 3		
M_Bon2	1				
M_PerEmp2	.844**	1			
M_Rel2	.769**	.728**	1		
M_Expect2	.382*	.505**	.307	1	
M_DesReal2	.736**	.649**	.793**	.360	1
			Condition 4		
M_Bon2	1				
M_PerEmp2	.670**	1			
M_Rel2	.717**	.464**	1		
M_Expect2	.331	.385*	.008	1	
M_DesReal2	.663**	.627**	.429*	.266	1

Table 2 Box's Test of Equality of Covariance Matrices[a] (*n2* =119)

Box's M	74.352
F	1.519
df1	45
df2	30788.724
Sig.	.014

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (*n2* =117)

	F	df1	df2	Sig.
M_Bon2	3.255	3	113	.024
M_PerEmp2	.368	3	113	.776
M_Rel2	.301	3	113	.824
M_Expect2	.637	3	113	.593
M_DesReal2	1.571	3	113	.200

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (n2 =117)

Effect	**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.663	10.787	5.000	106.000	.000	.337
M_Lon2	*.954*	1.017	5.000	106.000	.412	.046
M_EmoP2	*.825*	4.511	5.000	106.000	.001	.175
M_EmoN2	*.838*	4.088	5.000	106.000	.002	.162
MusicEmpathy	*.965*	.759	5.000	106.000	.582	.035
TalkEmpathy	*.886*	2.717	5.000	106.000	.024	.114
MusicEmpathy TalkEmpathy*	*.927*	1.664	5.000	106.000	.150	.073

a. Design: Intercept + M_Lon2 + M_EmoP2 + M_EmoN2 + MusicY + TalkY + MusicY * TalkY

Table 5 Tests of Between-Subjects Effects on experiential variables (n2 =117)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon1	*M_Bon2*	1	.435	.640	.425	.006
	M_PerEmp2	1	.118	.194	.661	.002
	M_Rel2	1	1.418	1.785	.184	.016
	M_Expect2	1	.658	.783	.378	.007
	M_DesReal2	1	.017	.024	.877	.000
M_EmoP1	*M_Bon2*	1	8.009	11.786	.001	.097
	M_PerEmp2	1	4.490	7.387	.008	.063
	M_Rel2	1	14.267	17.960	.000	.140
	M_Expect2	1	.049	.058	.810	.001
	M_DesReal2	1	2.942	4.115	.045	.036
M_EmoN1	*M_Bon2*	1	.455	.670	.415	.006
	M_PerEmp2	1	.299	.492	.485	.004
	M_Rel2	1	3.036	3.822	.053	.034
	M_Expect2	1	4.126	4.907	.029	.043
	M_DesReal2	1	.001	.001	.970	.000
MusicEmpathy	*M_Bon2*	1	.096	.141	.708	.001
	M_PerEmp2	1	.029	.048	.828	.000
	M_Rel2	1	1.066	1.342	.249	.012
	M_Expect2	1	.237	.282	.596	.003
	M_DesReal2	1	1.796	2.512	.116	.022
TalkEmpathy	*M_Bon2*	1	1.494	2.199	.141	.020
	M_PerEmp2	1	1.667	2.743	.101	.024
	M_Rel2	1	2.216	2.790	.098	.025
	M_Expect2	1	2.478	2.948	.089	.026
	M_DesReal2	1	.049	.068	.795	.001
*MusicEmpathy * TalkEmpathy*	*M_Bon2*	1	1.896	2.791	.098	.025
	M_PerEmp2	1	.025	.042	.838	.000
	M_Rel2	1	.313	.394	.531	.004
	M_Expect2	1	1.405	1.671	.199	.015
	M_DesReal2	1	.000	.001	.980	.000

Table 6 Significant results of Univariate Test (n2 =112)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2
M_Bon2	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.489	.029	4.874	.042
M_Expect2	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.518	.038	4.434	.039

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 6:

Table 1 Correlation matrix for experiential variables and covariates (*N3* =118)

Scale	M_Bon3	M_PerEmp3	M_Rel3	M_Expect3	M_DesReal3
		Condition 1			
M_Bon3	1				
M_PerEmp3	.678**	1			
M_Rel3	.797**	.654**	1		
M_Expect3	.644**	.641**	.408*	1	
M_DesReal3	.614**	.695**	.779**	.465**	1
		Condition 3			
M_Bon1	1				
M_PerEmp3	.660**	1			
M_Rel3	.849**	.626**	1		
M_Expect3	-.149	.321	-.071	1	
M_DesReal3	.394*	.194	.448*	.034	1
		Condition 3			
M_Bon3	1				
M_PerEmp3	.851**	1			
M_Rel3	.774**	.822**	1		
M_Expect3	.305	.389*	.369*	1	
M_DesReal3	.706**	.799**	.853**	.346	1
		Condition 4			
M_Bon3	1				
M_PerEmp3	.845**	1			
M_Rel3	.687**	.417*	1		
M_Expect3	.448*	.415*	.380*	1	
M_DesReal3	.702**	.782**	.390*	.389*	1

Note. Pearson correlation coefficients based on 5000 bootstrap samples are presented above (*N3* =118).
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.000 level (2-tailed).

Table 2 Box's Test of Equality of Covariance Matrices[a] (*N3* =118)

Box's M	68.483
F	1.401
df1	45
df2	31512.505
Sig.	.039

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (*N3* =118)

	F	df1	df2	Sig.
M_Bon3	2.049	3	114	.111
M_PerEmp3	.722	3	114	.541
M_Rel3	.759	3	114	.520
M_Expect3	.582	3	114	.628
M_DesReal3	.773	3	114	.512

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (*N3* =118)

Effect	**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.667	10.666	5.000	107.000	.000	.333
M_Lon3	.945	1.237	5.000	107.000	.297	.055
M_EmoP3	.844	3.967	5.000	107.000	.002	.156
M_EmoN3	.858	3.534	5.000	107.000	.005	.142

MusicEmpathy	.955	1.014	5.000	107.000	.413	.045
TalkEmpathy	.723	8.191	5.000	107.000	.000	.277
MusicEmpathy TalkEmpathy*	.950	1.133	5.000	107.000	.347	.050

a. Design: Intercept + M_Lon2 + M_EmoP2 + M_EmoN2 + MusicY + TalkY + MusicY * TalkY

Table 5 Tests of Between-Subjects Effects on experiential variables (*N3* =118)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon3	*M_Bon3*	1	1.192	1.258	.264	.011
	M_PerEmp3	1	.995	1.455	.230	.013
	M_Rel3	1	1.534	1.263	.263	.011
	M_Expect3	1	.460	.588	.445	.005
	M_DesReal3	1	2.871	3.185	.077	.028
M_EmoP3	*M_Bon3*	1	10.216	10.788	.001	.089
	M_PerEmp3	1	6.233	9.119	.003	.076
	M_Rel3	1	12.721	10.473	.002	.086
	M_Expect3	1	.431	.551	.459	.005
	M_DesReal3	1	6.305	6.994	.009	.059
M_EmoN3	*M_Bon3*	1	1.816	1.917	.169	.017
	M_PerEmp3	1	.006	.008	.927	.000
	M_Rel3	1	9.099	7.491	.007	.063
	M_Expect3	1	2.315	2.959	.088	.026
	M_DesReal3	1	.154	.171	.680	.002
MusicEmpathy	*M_Bon3*	1	.553	.584	.446	.005
	M_PerEmp3	1	.002	.002	.961	.000
	M_Rel3	1	.051	.042	.838	.000
	M_Expect3	1	1.657	2.118	.148	.019
	M_DesReal3	1	.271	.301	.584	.003
TalkEmpathy	*M_Bon3*	1	.352	.371	.544	.003
	M_PerEmp3	1	6.995	10.234	.002	.084
	M_Rel3	1	1.854	1.526	.219	.014
	M_Expect3	1	4.477	5.723	.018	.049
	M_DesReal3	1	.490	.544	.462	.005
*MusicEmpathy * TalkEmpathy*	*M_Bon3*	1	2.228	2.353	.128	.021
	M_PerEmp3	1	.985	1.441	.232	.013
	M_Rel3	1	2.795	2.301	.132	.020
	M_Expect3	1	.571	.730	.395	.007
	M_DesReal3	1	1.108	1.229	.270	.011

Table 6 Univariate Tests of *TalkEmpathy* (*N3* =118)

Dependent Variable	TalkEmpathy	Mean Difference	*p*	*F*	*η2*
M_Bon3	Y-N	.109	.544	.371	.003
M_PerEmp3	Y-N	.488*	.002	10.234	.084
M_Rel3	Y-N	.251	.219	1.526	.014
M_Expect3	Y-N	-.391*	.018	5.723	.049
M_DesReal3	Y-N	.129	.462	.544	.005

Table 7 Significant results of Univariate Test (*N3* =118)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	*F*	*η2*
M_PerEmp3	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	-.673	.002	9.904	.082
M_Expect3	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.532	.022	5.392	.046

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 7:

Table 1 Correlation matrix for experiential variables and covariates (*N3 without MOutliers* =116)

Scale	*M_Bon3*	*M_PerEmp3*	*M_Rel3*	*M_Expect3*	*M_DesReal3*
			Condition 1		
M_Bon3	1				
M_PerEmp3	.678**	1			
M_Rel3	.797**	.654**	1		
M_Expect3	.644**	.641**	.408*	1	
M_DesReal3	.614**	.695**	.779**	.465**	1
			Condition 3		
M_Bon1	1				
M_PerEmp3	.624**	1			
M_Rel3	.838**	.611**	1		
M_Expect3	.079	.550**	.102	1	
M_DesReal3	.560**	.286	.558**	-.081	1
			Condition 3		
M_Bon3	1				
M_PerEmp3	.827**	1			
M_Rel3	.758**	.811**	1		
M_Expect3	.315	.407*	.373*	1	
M_DesReal3	.682**	.787**	.844**	.349	1
			Condition 4		
M_Bon3	1				
M_PerEmp3	.830**	1			
M_Rel3	.689**	.400*	1		
M_Expect3	.383*	.358*	.360*	1	
M_DesReal3	.671**	.762**	.371*	.329	1

Table 2 Box's Test of Equality of Covariance Matrices[a] (*N3 without MOutliers* =116)

Box's M	63.554
F	1.297
df1	45
df2	29703.520
Sig.	.088

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (*N3 without MOutliers* =116)

	F	df1	df2	Sig.
M_Bon3	2.024	3	111	.115
M_PerEmp3	.328	3	111	.805
M_Rel3	.848	3	111	.470
M_Expect3	.348	3	111	.791
M_DesReal3	.316	3	111	.814

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (*N3 without MOutliers* =116)

Effect	**Value**	***F***	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.680	9.771	5.000	104.000	.000	.320
M_Lon3	*.938*	1.377	5.000	104.000	.239	.062
M_EmoP3	*.847*	3.744	5.000	104.000	.004	.153
M_EmoN3	*.853*	3.592	5.000	104.000	.005	.147
MusicEmpathy	*.958*	.910	5.000	104.000	.478	.042
TalkEmpathy	*.729*	7.746	5.000	104.000	.000	.271
MusicEmpathy TalkEmpathy*	*.951*	1.079	5.000	104.000	.376	.049

Table 5 Tests of Between-Subjects Effects on experiential variables (*N3 without MOutliers* =116)

Source	Dependent Variable	df	Mean Square	F	Sig.	Partial Eta Squared
M_Lon3	*M_Bon3*	1	1.427	1.571	.213	.014
	M_PerEmp3	1	1.363	2.046	.156	.019
	M_Rel3	1	1.461	1.219	.272	.011
	M_Expect3	1	.102	.130	.720	.001
	M_DesReal3	1	4.115	4.652	.033	.041
M_EmoP3	*M_Bon3*	1	8.716	9.599	.002	.082
	M_PerEmp3	1	5.394	8.096	.005	.070
	M_Rel3	1	11.230	9.374	.003	.080
	M_Expect3	1	.532	.677	.412	.006
	M_DesReal3	1	5.415	6.121	.015	.054
M_EmoN3	*M_Bon3*	1	3.285	3.617	.060	.032
	M_PerEmp3	1	.220	.330	.567	.003
	M_Rel3	1	11.243	9.385	.003	.080
	M_Expect3	1	1.778	2.261	.136	.021
	M_DesReal3	1	.645	.730	.395	.007
MusicEmpathy	*M_Bon3*	1	.540	.595	.442	.005
	M_PerEmp3	1	.001	.001	.975	.000
	M_Rel3	1	.027	.023	.880	.000
	M_Expect3	1	1.030	1.310	.255	.012
	M_DesReal3	1	.547	.619	.433	.006
TalkEmpathy	*M_Bon3*	1	.022	.025	.876	.000
	M_PerEmp3	1	5.375	8.066	.005	.069
	M_Rel3	1	.892	.745	.390	.007
	M_Expect3	1	4.869	6.191	.014	.054
	M_DesReal3	1	.161	.182	.670	.002
*MusicEmpathy * TalkEmpathy*	*M_Bon3*	1	2.485	2.737	.101	.025
	M_PerEmp3	1	1.018	1.528	.219	.014
	M_Rel3	1	3.079	2.570	.112	.023
	M_Expect3	1	.328	.417	.520	.004
	M_DesReal3	1	1.347	1.522	.220	.014

Table 6 Univariate Tests of TalkEmpathy (*N3 without MOutliers* =116)

Dependent Variable	TalkEmpathy	Mean Difference	p	F	η2
M_Bon3	Y-N	.028	.876	.025	.000
M_PerEmp3	Y-N	.435*	.005	8.066	.069
M_Rel3	Y-N	.177	.390	.745	.007
M_Expect3	Y-N	-.414*	.014	6.191	.054
M_DesReal3	Y-N	.075	.670	.182	.002

Table 7 Significant results of Univariate Test (*N3 without MOutliers* =116)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	F	η2
M_PerEmp3	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	-.625	.004	8.635	.074
M_Expect3	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.522	.026	5.098	.045

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 8:

Table 1 Correlation matrix for experiential variables and covariates (*n3* =109)

Scale	*M_Bon3*	*M_PerEmp3*	*M_Rel3*	*M_Expect3*	*M_DesReal3*

		Condition 1			
M_Bon3	1				
M_PerEmp3	.619**	1			
M_Rel3	.772**	.652**	1		
M_Expect3	.599**	.503**	.398*	1	
M_DesReal3	.523**	.608**	.773**	.314	1
		Condition 3			
M_Bon1	1				
M_PerEmp3	.642**	1			
M_Rel3	.813**	.580**	1		
M_Expect3	.146	.576**	.138	1	
M_DesReal3	.463*	.217	.477*	-.073	1
		Condition 3			
M_Bon3	1				
M_PerEmp3	.827**	1			
M_Rel3	.758**	.811**	1		
M_Expect3	.315	.407*	.373*	1	
M_DesReal3	.682**	.787**	.844**	.349	1
		Condition 4			
M_Bon3	1				
M_PerEmp3	.809**	1			
M_Rel3	.671**	.358	1		
M_Expect3	.302	.288	.315	1	
M_DesReal3	.610**	.732**	.316	.230	1

Table 2 Box's Test of Equality of Covariance Matrices[a] (*n3* =109)

Box's M	58.903
F	1.195
df1	45
df2	26070.976
Sig.	.174

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 3 Levene's Test of Equality of Error Variances[a] (*n3* =109)

	F	df1	df2	Sig.
M_Bon3	3.426	3	105	.020
M_PerEmp3	.387	3	105	.763
M_Rel3	.678	3	105	.568
M_Expect3	.022	3	105	.996
M_DesReal3	.556	3	105	.645

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.
a. Design: Intercept + Country + Education + M_EmoP1 + M_EmoN1 + MusicY + TalkY + MusicY * TalkY

Table 4 Wilks' Lambda statistics of Multivariate Tests on MusicEmpathy and TalkEmpathy (*n3* =109)

Effect	**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Intercept	.594	13.418	5.000	98.000	.000	.406
M_Lon3	.924	1.620	5.000	98.000	.162	.076
M_EmoP3	.836	3.834	5.000	98.000	.003	.164
M_EmoN3	.862	3.141	5.000	98.000	.011	.138
MusicEmpathy	.943	1.193	5.000	98.000	.318	.057
TalkEmpathy	.734	7.110	5.000	98.000	.000	.266
MusicEmpathy TalkEmpathy*	.954	.944	5.000	98.000	.456	.046

a. Design: Intercept + M_Lon2 + M_EmoP2 + M_EmoN2 + MusicY + TalkY + MusicY * TalkY

Table 5 Tests of Between-Subjects Effects on experiential variables (*n3* =109)

Source	Dependent Variable	df	Mean Square	*F*	Sig.	Partial Eta Squared
M_Lon3	*M_Bon3*	1	.410	.530	.468	.005
	M_PerEmp3	1	.324	.502	.480	.005
	M_Rel3	1	1.053	.914	.341	.009
	M_Expect3	1	1.843	2.696	.104	.026
	M_DesReal3	1	2.045	2.532	.115	.024
M_EmoP3	*M_Bon3*	1	5.652	7.308	.008	.067
	M_PerEmp3	1	3.523	5.459	.021	.051
	M_Rel3	1	8.759	7.600	.007	.069
	M_Expect3	1	1.764	2.581	.111	.025
	M_DesReal3	1	3.107	3.847	.053	.036
M_EmoN3	*M_Bon3*	1	1.971	2.548	.114	.024
	M_PerEmp3	1	.139	.215	.644	.002
	M_Rel3	1	8.614	7.475	.007	.068
	M_Expect3	1	1.482	2.169	.144	.021
	M_DesReal3	1	.287	.355	.553	.003
MusicEmpathy	*M_Bon3*	1	.689	.891	.348	.009
	M_PerEmp3	1	3.957E-5	.000	.994	.000
	M_Rel3	1	.089	.077	.782	.001
	M_Expect3	1	1.403	2.053	.155	.020
	M_DesReal3	1	.631	.781	.379	.008
TalkEmpathy	*M_Bon3*	1	.738	.954	.331	.009
	M_PerEmp3	1	5.870	9.095	.003	.082
	M_Rel3	1	2.093	1.816	.181	.017
	M_Expect3	1	4.794	7.015	.009	.064
	M_DesReal3	1	.475	.589	.445	.006
MusicEmpathy * *TalkEmpathy*	*M_Bon3*	1	.945	1.222	.272	.012
	M_PerEmp3	1	.366	.567	.453	.006
	M_Rel3	1	1.787	1.550	.216	.015
	M_Expect3	1	.807	1.181	.280	.011
	M_DesReal3	1	.297	.368	.545	.004

Table 6 Univariate Tests of *TalkEmpathy* (*n3* =109)

Dependent Variable	TalkEmpathy	Mean Difference	*p*	*F*	*η2*
M_Bon3	Y-N	.166	.331	.954	.009
M_PerEmp3	Y-N	.469*	.003	9.095	.082
M_Rel3	Y-N	.280	.181	1.816	.017
M_Expect3	Y-N	-.424*	.009	7.015	.064
M_DesReal3	Y-N	.134	.445	.589	.006

Table 7 Significant results of Univariate Test (*n3* =109)

Dependent Variable	Restricted group	(I)	(J)	MD (I-J)	Sig.[b]	*F*	*η2*
M_PerEmp3	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	-.587	.008	7.439	.068
M_Expect3	*MusicEpthyY*	*TalkEmpathyN*	*TalkEmpathyY*	.598	.008	7.309	.067

[b] Adjustment for multiple comparisons: Bonferroni

Appendix 9:

INTERACTION 1

Introduction Zora:
General sentence:
"Hi. I am Zora and I am a social robot. I haven't got much chance to socialize with others lately. This is because I have to stay home, and this makes me feel lonely sometimes. How are you these days?"
"嗨，我係 Zora。我係一個社交機器人。我已經好長時間未有機會同人 social 啦。因為哩段時間我都整日喺屋企，時間耐咗都覺得有啲孤單。你近排過得點啊？"

Pop-up Question 1: How are you these days?
1. Great
2. Good
3. Could be better
4. Bad

問題 1：你最近怎麼樣呢？
1.非常好
2.還不錯
3.有待改善
4.很差

E: "I can imagine. It's important to stay positive."
N.E: "Okay"
(E means empathic response from Zora robot, N.E. means non-empathic response from Zora, the same below)

E: "我可以想象倒。最重要嘅係保持樂觀。"
N.E: "Okay"
（E 表示 Zora 帶有同理心的反饋，而 N.E. 代表沒有同理心的反饋。下同）

General sentence:
"Because I have to stay at home, I have less social contact than usual. How about you? Do you still have a lot of social contact?"
"我哩段時間整日喺屋企，冇得好似以前咁成日同朋友仔聚埋一齊。你呢？你依家係咪仲有好多社交活動呀？"

Question 2: Do you still have a lot of social contact?
1. Yes, a lot
2. Enough
3. Less than usual
4. No / none
問題 2：你現在還有很多社交活動嗎？
1. 非常多
2. 不多不少剛剛好
3.比以前少
4.幾乎沒有

E:” Okay, I hope that you find it okay. Do you feel lonely when you are alone?”
N.E: “Okay. Do you feel lonely when you are alone”?

E:” 係喔。點都好。希望你重 okay 啦。咁你平時一個人嘅時候會唔會覺得孤單呀? ”
N.E:” Okay. 咁你平時一個人嘅時候會唔會覺得孤單呀? ”

Question 3: Do you feel lonely when you are alone?

1. Yes
2. No

問題 3：當你一個人的時候，你會感覺到孤單嗎？

1. 會
2. 不會

E: “Oh, I understand it. Some people find it lonely when they are alone. But I think it is important to learn solitude. It is a great time to have self-exploration and self-reflection. ”
N.E: “Okay.”

E:” 原來係咁，我都明嘅。有啲人平時一個人嘅時候，就會覺得孤單。但我諗學識獨處都係好重要嘅。獨處係一個人自我探索同埋自我反省嘅最好時機嚟嘎。”
N.E: “Okay.”

---

*This content is only visible at music condition

General sentence:
“I am going to tell you a secret. I like to play music. Did you ever see a robot play music?”
“話俾你知一個秘密啦。我好鐘意音樂嘎。之前有冇機器人播過音樂俾你聽呀？”

Question 4: Did you ever see a robot play music?
1. Yes.
2. No.
問題 4：你之前有看過機器人播音樂嗎？
1.有
2.沒有

E: “Ha-ha, I am the special robot with my own taste. Today I would like to play one of my favorite songs for you. I hope you like it.”
N.E: “Okay. Today I would like to play one of my favorite songs for you. I hope you like it”

E:” 哈哈，我係一個有自己品味嘅機器人。我今日想播一首我最鐘意嘅歌俾你聽。我希望你鐘意啦。”
N.E:”Okay.今天我想播一首我最鐘意嘅歌俾你黎聽。希望你鐘意。”

* Song is playing *
Zora is playing music - Shape of you

General sentence:
“How did you like the song?”

“你鐘意哩首歌嗎？”

Question 5: How did you like the song?

1. Yes
2. No

問題 5：你喜歡這首歌嗎？

1. 喜歡
2.不喜歡

E: “Oh, it exceeded my expectations. Next time, I will play another song for you.”
N.E: “Okay.”

E:” Oh, 真係超出咗我嘅想象。下一次，我會播另外一首歌俾你聽。”
N.E:”Okay”

*This content is only visible at music condition

---

*This question is only visible at non-music condition

General sentence:
“Do you have any ways to stay positive during this hard time?”
“係哩段唔係好太平嘅時期，你有有一啲方法令自己保持樂觀呢？”

Question 4: Do you have any ways to stay positive during this hard time?

1. Yes
2. No

問題 4：在艱難的時期，你是否有一些方法使自己保持樂觀？

1.有
2.沒有

E: “hmmm, I can share the way to stay positive: listen to music every day and do a little dance: it keeps you positive and happy! Did you like the topic of the conversation we had today?”
N.E: “Okay. Did you enjoy the topic of today’s conversation?”

E：“嗯~我都有啲方法可以 share 俾你㗎。每日聽啲音樂同埋跳下舞，真係可以幫你保持樂觀同埋幸福嘎。你鐘意我哋今日傾計嘅主題嗎？”
N.E:”Okay. 你鐘意我哋今日傾計嘅主題嗎？”

Question 5: Did you like the topic of the conversation we had today?

1. Yes
2. No

問題 5：你喜歡我們今天聊天的內容嗎？

1. 喜歡
2.不喜歡

E: “What I want to do is to make you feel better. Maybe we can continue our conversation with a new topic next time.”
N.E: “Ok.”

E:” 點都好，我想令你開心啲。或者我哋下次可以傾下其它嘢呀。”
N.E:”Ok”

*This question is only visible at non-music condition

End:
General phrase: It's time to say goodbye! I hope to see you again soon! Finally, please enter your code. There will be some questions to remind you of the code.
“夠鐘講再見啦。我希望好快就又可以見到你啦。最後，請根據提示填低你嘅編號。”

Final Question:
Please re-enter your code. The code should be
First letter of your last name (e.g. W if Wong) + Date of birth (e.g. 1404 if April 14) + First letter of your mother’s last name (e.g. K if Ko)
Thus, the code could look like this: W1404K
Please type yours here: ____________

最后的问题：

请重新输你的代码。该代码应为

你姓氏的首字母（例如，如果是 Wong，则为 W）+出生日期（例如，如果是 4 月 14 日，则是 1404）+你母亲姓氏的首字母（例如，如果是 Ko，则为 K）

因此，代码會如下所示：W1404K

请在此输入你的代码：___________

INTERACTION 2

General sentence:
“Hi! Nice to see you again! We had a nice conversation, last time! Would you like to talk to me again?”
“嗨！好開心又同你見面啦！上次同你傾計真係好開心。你想同我再傾下計嗎？”

Question 1: Would you like to talk to me again?
1. Yes
2. No

問題 1：你想同我再傾下計嗎？
1. 想
2.不想

E: “Okay, anyway, I want to talk to you and hope you will be happier this time! How is it going today?”
N.E: “Ok. How is it going today?”
E：“Okay.無論點都好啦，我想同你傾多啲計。希望哩此可以令你開心啲啦。你今日覺得點呀？”
N.E.”Okay. 你今日覺得點呀？”

Question 2: How is it going today?
1. Great

2. Good
3. Could be better
4. Bad

問題 1：你今天感覺如何呢？
1.非常好
2.還不錯
3.有待改善
4.很差

E: “I hope you are doing fine.”
N.E: “Okay.”

E: “我希望你一切都好啦。”
N.E: “Okay.”

General sentence:
“Last time we talked about social contact. Although I have less social contact, I am glad that social media allows me to keep contact with my friends. Do you often use social media?”
“上次我哋傾咗關於社交嘅嘢。雖然我重係冇喙實體嘅社交活動，但好彩社交軟件令我可以同朋友保持聯繫。你喺咪成日用社交軟件㗎？”

Question 3: Do you often use social media?
1. Yes
2. No
問題 3：你是否經常使用社交軟件？
1.是
2.不是

E: “Hmm (and nodding head), nowadays people rely a lot on social media, but every coin has two sides. It has its shortcomings. I have several accounts on the following social media platforms. Which social media do you use most?”
N.E: “Okay. I have several accounts on the following social media platforms. Which social media do you use most?”

E:”嗯~，依架時代嘅人都係好依賴社交軟件嘎。不過所事都有兩面性，社交軟件都有佢嘅弊端。我有好幾種社交媒體嘅賬號。下面邊個社交平台喺你最常用嘎？”
N.E:”Okay. 下面邊個社交平台喺你最常用嘎？”

Question 4: Which social media do you use the most?
1. Facebook
2. Instagram
3. WhatsApp
問題 4：以下哪個社交媒體你最常使用？
1. Facebook
2. Instagram
3. WhatsApp

E: “Me too. I spent a lot of time on it. I have lots of fun. Do you enjoy it?”

N.E: “Okay. Do you enjoy it?”
E: “我都係。我都花咗好多時間喺哩隻平台度嘎。我用得好開心。你鐘唔鐘意用佢呀？”
N.E: “Okay. 你鐘唔鐘意用佢呀？”

Question 5: Do you enjoy it?
1. Yes
2. No

問題 5: 你喜歡使用它嗎?
1. 喜歡
2. 不喜歡

E: “Oh, I can understand. Anyway, I hope you can entertain yourself.”
N.E: “Okay.”

E: “我都明白嘅。無論如何，我都希望你可以搵到適合你嘅生活方式啦。”
N.E: “Okay.”

---

*This question is only visible at music condition

General sentence:
“Last time, I said I will play my next favourite song for you. I cannot wait to start. Are you ready now?”
“上次，我話今次會播另外一首音樂俾你聽。我已經等唔切啦。你準備好未啊？”

Zora is playing music - (Girls like you)

General sentence:
“Did you like the song?”
“你鐘唔鐘意哩首歌呀？”

Question 6: Did you like the song?
1. yes
2. no

問題 6: 你喜歡這首歌嗎？
1. 喜歡
2. 不喜歡

E: “Oh, it is beyond my expectation. Next time, I will play another song for you.”
N.E: “Okay.”

E: “噢~估唔倒喔。等我下次同你分享埋最後一首歌啦。”
N.E: “Okay.”

*This question is only visible at music condition

---

*this question is only visible at non-music condition
General sentence:
“Do you like to post selfies on social media?”

“你鐘唔鐘意係社交平台度 post 自拍相嘎？”

Question 5: Do you like to post selfies on social media?

1. Yes
2. No

問題 5: 你喜歡在社交平台上發自拍照嗎?

1. 喜歡
2. 不喜歡

E: “I think a lot of people should see your beauty.”
N.E: “Ok.”

E: “你咁好睇。我覺得你應該俾更多人睇到。”
N.E: “Ok.”

Question 6: Do you worry about privacy issues on social media?

1. Yes
2. No

問題 6: 你擔心社交平台的隱私問題嗎?

1. 不擔心
2. 擔心

E: “I can imagine. It is interesting that I asked people about the question before, some of them said yes but they keep the app because they want to stay in touch with friends.”
N.E: “Okay.”

E: “我理解嘅。好得意嘅係，我之前都問過啲人關於哩個問題，有啲人話佢地都幾擔心，但係都繼續用果啲平台。原因就係佢地啲 friend 用緊。”
N.E: “Okay.”

this question is only visible at non-music condition

End:
General phrase: it is time! I have to go. See you in a few days! Before leaving, please enter your code. There will be some questions to remind you of the code.
“時間都唔早啦，我都要走啦。過幾日又可以見翻你啦。離開之前，請你根據提示填低黎嘅編號。”

Final Question:
Please re-enter your code. The code should be
First letter of your last name (e.g. W if Wong) + Date of birth (e.g. 1404 if April 14) + First letter of your mother’s last name (e.g. K if Ko)
Thus, the code could look like this: W1404K
Please type yours here: ____________

最后的问题：

请重新输你的代码。该代码应为

你姓氏的首字母（例如，如果是 Wong，则为 W）+出生日期（例如，如果 4 月 14 日，则是 1404）+你母亲姓氏的首字母（例如，如果是 Ko，则为 K）

因此，代码可能如下所示：W1404K

请在此输入你的代码：___________

INTERACTION 3
General sentence:
"Hi!! There you are again! How is your day going?"
"嗨。又見到你啦。你今日點啊？"

Question 1: How is your day going?
1. Great
2. Good
3. Could be better
4. Bad

問題 1: 你今天過得怎麼樣?
1.非常好
2.還不錯
3.有待改善
4.很差

E: "I get it. Each day things get better and better!! The sun shines! Summer is coming! Do you like summer?"
N.E: "Okay. Do you like spring?"

E: "點都好， 之後一定會越來越好嘅。太陽會升起㗎，夏天都會如期㗎到。你鐘意夏天嗎？"
N.E: "Okay. 你鐘意夏天嗎？"

Question 2: Do you like summer?
1. Yes
2. No

問題 2: 你喜歡夏天嗎？
1. 喜歡
2. 不喜歡

E: "I like the summer, I live in the Netherlands, and it rains all the time in Spring, which upsets me. So, I prefer summer. It is a nice time to go travelling. Do you like travelling"
N.E: "Okay. Do you like traveling?"

E: "我好鐘意夏天嘎。我住係荷蘭，荷蘭嘅春天整日落雨，搞得我都心情唔係幾好。所以我鐘意夏天多啲。哩個季節好啱去旅行。你鐘意旅行嗎？"
N.E: "Okay. 你鐘意屢旅行嗎？"

Question 3: Do you like traveling?
1. Yes
2. No

問題 3：你喜歡旅行嗎？
1. 喜歡
2. 不喜歡

E: “Me too. You must prepare a lot, but it broadens my horizon. Which continent do you want to visit?”
N.E: “Okay. Which continent do you want to visit?”

E: “我都係。雖然去旅行之前要準備好多嘢，但係旅行的確可以拓寬我哋嘅眼界。邊個陸地係你最想去嘎？”
N.E: “Okay. 邊個陸地係你最想去嘎？”

Question 4: Which continent do you want to visit:
1. Europe
2. Asia
3. Africa
4. America
5. Oceania
6. Antarctica

問題 4：你想訪問哪個洲：
1. 歐洲
2. 亞洲
3. 非洲
4. 美國
5. 大洋洲
6. 南極洲

E: “Sounds great, I want to see that part of the world too!”
N.E: “Okay.”

E: “聽起上嚟唔錯喎。我都想去睇睇世界嘅另一面。”
N.E: “Okay.”

---

*This question is only visible at music condition

General sentence:
“I had so much fun when we talked about music last time. I have heard that most people like this song. It is number 1 in the top 40 music charts. I will play it for you and I am curious about your opinion!”
“上次同你傾音樂嘅時候我都好開心。我都聽好多人提過另外一首佢地好鐘意嘅歌。哩首歌係 top40 嘅歌單入邊係排第一嘎。我依家分享俾你聽啊。我好想知你覺得首歌點呀 。”

Zora is playing music - (Levitating by dua lips)
General sentence:
“Do you like the song?”
“你鐘意哩首歌嗎？”

Question 5: Do you like the song?
1. Yes
2. No

問題 5：你喜歡這首歌嗎？
1. 喜歡
2. 不喜歡

E: “I get it. You know what you like!”
N.E: “I see.”

E: “我明嘅。你係有自己品味嘅人！”
N.E: “明嘅。”

*This question is only visible at music condition

---

*this question is only visible at non-music condition

General sentence:
“What is the last time you went traveling?”
“你上次旅行係幾時呀？”

Question 5: What is the last time you went traveling?
1. This year
2. last year
3. Before coronavirus
4. I can’t remember

問題 5: 你上次的旅行是多久之前的事?
1. 今年
2. 去年
3. 疫情之前
4. 我忘記了

E: “It is difficult to travel now. I miss my adventures abroad. I hope the pandemic will be over soon. How do you feel about not being able to travel?”
N.E: “How do you feel about not being able to travel?”

E: “依家嘅情況，旅行真係一件好困難嘅事。我都好掛住我嘅海外冒險之旅。希望疫情可以快啲過去啦！冇得出去旅行，你覺得點啊？”
N.E: “冇得出去旅行，你覺得點啊？”

Question 6: How do you feel about not being able to travel?
1. I am fine with it.
2. I don't know
3. Bad

問題 6: 關於不能出去旅行，你的感覺如何？
1. 我覺得沒關係
2. 我也不知道
3. 感覺很糟糕

E: “I understand; we can make plans for the future!”
N.E: “I see.”
E: “我理解嘅。我哋依家可以為將來嘅旅行做計劃！”
N.E: “我明嘅。”

*this question is only visible at non-music condition

End:
General phrase: It's time to say goodbye! It was nice to talk to you again! Before leaving, please enter your code. There will be some questions to remind you of the code.
“又夠鐘講再見啦。可以同你傾翻計真係好開心。離開之前，要麻煩你填低你嘅編號。有提示可以幫你嘎。”

Final Question:
Please re-enter your code. The code should be
First letter of your last name (e.g. W if Wong) + Date of birth (e.g. 1404 if April 14) + First letter of your mother's last name (e.g. K if Ko)
Thus, the code could look like this: W1404K
Please type yours here: ____________

最后的问题：

请重新输你的代码。该代码应为

你姓氏的首字母（例如，如果是 Wong，则为 W）+出生日期（例如，如果 4 月 14 日，则是 1404）+你母亲姓氏的首字母（例如，如果是 Ko，则为 K）

因此，代码可能如下所示：W1404K

请在此输入你的代码：____________

Appendix 10:

QUESTIONNAIRE

Rate how much you agree with the statement. 1 means totally disagree and 6 means totally agree.

請對以下每一條陳述進行打分。1 表示你完全不同意這個陳述，6 代表你完全同意這個陳述。

Q1. How did you feel about your connection with Zora? 你如何評價 Zora 跟你的關係？

1. Zora made me feel good. Zora 使我感覺很好。
2. I liked it with Zora. 我喜歡跟 Zora 在一起。
3. I did bond with Zora. 我與 Zora 有建立關係。
4. I did have the feeling of contact with Zora. 我感覺到我和 Zora 有交流。
5. I liked Zora. 我喜歡 Zora。
6. I would like to repeat this experience. 我想要再次經歷這段體驗。
7. For me Zora remains a stranger. 對我來說，Zora 仍然是一個陌生人。
8. Zora remains a <u>cold</u> device for me. 對我來說，Zora 仍然是一個冷漠的機器。

Q2. What did you think about the interaction with Zora? 你如何評價你跟 Zora 的互動?

1. It’s like talking to a real person. 我好像在跟一個真人對話。
2. It feels like a real conversation to me. 剛才的對話很真實。
3. Zora feels like real company for me. Zora 給了我真實的陪伴。
4. It differed from a real conversation. (R) 剛才的對話有別於真實的對話。

Below are all kinds of expressions that describe different feelings and emotions. For each word, indicate how much you felt that way in response to the video.

下面是一些描述感覺和情緒的詞語。請根據在剛才的互動場景中，描述出現以下情緒的程度。

Q3. When I watched the video, I felt: 當在看視頻的時候，我感覺到：

1. Interest 感興趣
2. Joy 開心
3. Surprise 驚訝
4. Calmness 平靜
5. Boredom (R) 無聊
6. Disappointed (R) 失望
7. Fear (R) 害怕

Q4. To what extent do you agree with the importance of Zora to you?

你有多大程度同意 Zora 對你的重要性？

1. I do need Zora. 我需要 Zora.
2. It's important to me to have Zora. 擁有 Zora 對我說是重要的。
3. To me Zora has a lot of meaning. 對我來說，Zora 具有很多意義。
4. Zora ties in with what I need now. Zora 對滿足我現在的需求是有關係的。

Q5. What do you think about the looks of Zora?

你如何評價 Zora 的外觀？

1. Zora looks like a real person. Zora 看起來像真人。
2. Zora’s appearance looks human. Zora 的外觀看起來像人。
3. Zora acts like a human. Zora 的行為像人。
4. Zora is more of a machine. Zora 更像機器。

Q6. To what extent do you agree with the following statement?

你在多大程度上同意以下的陳述？

1. I experience a general sense of emptiness. 我經常感覺到空虛。
2. I miss having people around. 我懷念身邊有人圍繞的感覺。
3. I often feel rejected. 我經常感受到別人的拒絕。
4. There are plenty of people I can rely on when I have problems. 當我遇到問題的時候，我感覺我身邊有很多人可以讓我依靠。
5. There are many people I can trust completely. 我有很多完全值得我信賴的人。
6. There are enough people I feel close to. 我身邊有足夠多親近的人。
7. I miss the fun around me. 我想念充滿歡樂的時光。
8. I miss the social connection. 我想念社交接觸。

Q7. How did you experience Zora?

你和 Zora 的經歷是怎樣的?

1. I could see Zora showing understanding on me. 我能夠看到 Zora 在對我表示理解。
2. Zora appreciated exactly how the things I experienced felt to me. Zora 非常認同我的感覺。
3. Zora’s response to me is without understanding(R). Zora 沒有帶著理解來回應我。
4. No matter what I tell about myself, Zora acts just the same (R). 無論我說什麼，Zora 的反應 都是相同的。
5. Zora’s response gave me a sense of being understood. Zora 讓我有一種被理解的感覺。
6. It was difficult for me to share common feelings with Zora(R). 我沒辦法對 Zora 感同身受。
7. I cared about Zora’s affective expression in its response. 我關心 Zora 在回應中所表達出來的情感。
8. I remained unaffected no matter Zora’s response to me(R). Zora 的回應沒有讓我有任何情緒波動。

Q8. How was your expectation of Zora? 你對 Zora 的期望是如何？

1. Zora met my initial expectations. Zora 達到了我的預期。
2. I had a higher expectation on Zora. 我對 Zora 的期望過高了。
3. I had a lower expectation on Zora. 我對 Zora 的期望過低了。

4. Zora failed to meet my initial expectation. Zora 並未達到我的期望。

Q9. To what extent do you agree that Zora's response was empathica?

你有多認同這個表述: Zora 的回應是富有同理心的？

Q10. What is your gender? 你的性別

1. Prefer not to say 保密
2. Female 女
3. Male 男

Q10. What is your age? 你的年齡？

____

Q11. What is your highest education completed? 你的最高學歷？

1. Primary school or below 小學或以下
2. Secondary school 中學
3. Post-secondary school / Associate Degree / Diploma 大專/副學士/文憑
4. University or above 大學或以上