

Emotionality in Computer-Mediated Environments: Analyzing the Emotional Content of Psychotherapeutic Communication

Ebrahim Oshni Alvandi, Monash University and Royal Melbourne Institute of Technology

Geroge Van Doorn, Federation University, Australia

Mark Symmons, Monash University

The present study aimed to investigate the ability of psychotherapists to decipher emotional discourse in computer-mediated communication and whether delivery mediums can increase the ambiguity in their assessment. Ignoring any reduction in efficacy of remote therapy, it is assumed that therapists' experiences of communication and their ability to construct a diagnosis on the basis of what they perceive will be influenced by Computer-Mediated Psychotherapy. Participants of the study scaled four emotional factors i.e., intensity, valence, potency and activation to measure the constructive nature of emotional discourse. Among other findings, the study showed that estimating emotional factors of communication did not differ significantly across main delivery modalities (i.e., text, audio and video). The results demonstrate further that telecommunication technologies did not alter the diagnostic level of psychotherapists. However, the reduction of emotional perception is very much remarkable for clinicians' decisions when discourse is communicated via non-human computer-mediated mode.

Key Words: Telecounseling, Cyberpsychotherapy, Human-Computer Interaction, Emotional Perception, Diagnosis

DOI: 10.24982/jois.1711017.002

1 INTRODUCTION

The ability to perceive clients' emotions¹, and to recognize the nonverbal indicators of emotions, has significant importance in therapy [1]. In Emotion Focused Therapy (EFT), for instance, attending to bodily reactions such as sighs, posture, vocal quality and facial expression can assist clinicians in assessing the maladaptive or destructive behaviors caused by defensive emotions [2, 3]. Moreover, it has been argued that psychotherapists who are better at recognizing behavioral cues and emotional expressions will have more satisfied patients and can provide more effective support [4].

While the behavioral cues available in face-to-face therapeutic communication contribute significantly to the perception of emotions (see [5], [6], and [7]); it may be that the perception of emotions in computer-mediated environments is tantamount to face-to-face settings. In support of this proposition, some studies have indicated that emotions are communicated as effectively via computer-mediated communication (CMC) as they are in face-to-face interventions. For example, early research from the field of human computer interactions (HCI) demonstrated that a user's online presence and interpersonal relationships are positively influenced by behavioral cues [8, 9]. Furthermore, it has recently been shown that users who use visual cues in an online environment are deemed more attractive [10]. Regardless of the reason to be seen or see someone appealing, this finding would indicate that considerable emotional information can be transmitted online using verbal cues and facial expressions, as well as emoticons [11].

¹Emotions, simply, are bio-physiological, cognitive changes or reactions caused by an imbalance between the object, goals, concerns, and the needs of individuals [12, 13].

Contrary to this possibility is the view that the transfer of emotional information in CMC is less clear than it is in face-to-face settings [14, 15]. It has been argued that emotional perception will be depleted due to (1) the intensity of experiencing emotional states, (2) the restrictions placed on interpreting verbal cues, and (3) the transmission of relevant, nonverbal characteristics [4, 16]. Some studies maintain that CMC reduces the amount and type of such information [17, 18]. In particular, computer-mediated technologies, such as instant messaging, lack full engagement between communicators [19, 20 and 21]. Also, behavioral cues such as body movements, which are important in the interpretation of emotions, are absent from audio transmission, or diminished in video-based communication [4, 21, and 22].

Clarifying the nature of emotional communication during CMP sessions has not become an important topic of enquiry. To the authors' knowledge, only one study has been undertaken to date that comprehensively explored such communication. In that study, Machado, Beutler, and Greenberg used stimuli from real psychotherapy sessions to test psychotherapists' ability to recognize emotions [23]. In their experiment, a therapeutic conversation was presented to participants either on video-tape or via a written, verbatim transcript and psychotherapists had to identify the client's expressed emotions (i.e., love, joy, surprise, anger, sadness, or fear). The ability to identify emotions was not significantly altered by the intensity with which the emotion was expressed, or by videotape versus transcript modes of delivery. Furthermore, there were no significant differences in the abilities of psychotherapists and people without training in psychology in recognizing emotions.

1.1 The Current Study

From the clinical perspective, emotional perception is the psychotherapist's proficiency in identifying emotional disorders. Psychotherapists have to be skilled at decoding client emotions [24]. This skill is significant because good emotional perception leads to a successful cognitive and emotional bond with clients, thereby positively orienting the therapy, exploring and reflecting on the clients' concerns, and focusing attention on their feelings. However, there is a paucity of literature assessing how emotional perception is applied during remote therapy. The current study uses a qualitative methodology and incorporates the perspectives on how clinicians perceive the emotional content presented via various modes of pre-recorded communication. While there are conflicting views on the ability of CMC to display behavioral cues; artificial audio (i.e., robotic voice), audio-only, video-only, and video-audio presentations of communication are employed to systematically determine which components of communication are important in assessing the emotional content.

Another question of interest in this study relates to whether emotional perception affects diagnoses made by psychotherapists. Clinicians determine the presence of a mental disorder based on analytical or heuristic evaluations of clients' emotional communication [25, 26]. The importance of heuristic decision-making has been noted in medicine, with some suggesting that less information causes less error and thus more accurate diagnoses [27]. This implies that more information gained from a client may not always benefit the clinician's decision-making. However, bodily and facial cues, as well as, audio information, are known to be critical for effective psychotherapy [16], and enhance the recognition of emotions [28]. This issue is significant when considering psychotherapy via technology because we must consider how the environment contributes to, or impedes, decision-making. Therefore, exploring the effects of the environment on one's ability to perceive emotions are necessary, as is emotional perception and decision-making during CMP.

One study has examined the decision-making ability of psychotherapists in relation to the level of emotional information transmitted via technology. Bambling, King, Reid, and Wegner showed that when emotions were displayed at a lower level of intensity, psychotherapists underestimated the severity of the client's problems during a text-based, synchronous, single-session of counseling

[29]. Given this information, the current study sought to examine the relationship between the *emotionality* of therapeutic communication and *decision-making* when communication technologies are employed. It seems non-controversial to say that if clinicians have access to, and perceive, differing levels of emotional information, they may not reach consistent conclusions with regard to a unified evaluation of the clients' intentions/issues.

1.1.1 Hypothesis

Despite the positive findings on emotional perception that CMC studies demonstrate, the challenge remains as to whether telecommunication technologies can provide adequate information for the clinician to perceive emotions accurately. Further investigation is required into how well CMC conveys emotionality, and how useful this information is for making decisions about the psychological problems of clients. Based on the literature:

- 1) The ability to perceive emotional content will be determined by the amount of, or more correctly number of, channels of, communication such that text-based communication will result in the worst performance compared to verbal (audio) communication which, in turn, will be hindered compared to video-audio communication.
- 2) It is reasonable to expect that the heuristic decision-making of clinicians will be enhanced as the availability of emotional information in communication increases (e.g., when verbal [audio] and visual [video] cues are introduced).

2 METHOD

2.1 Design and Analysis

The data were collected in two experimental conditions (emotion-laden and emotion-based scenarios) via five modes (text, robotic voice, audio-only, video-only, video-audio). For mixed-model ANOVA designs, "expertise in psychology" of participants was the between-subjects factor while, emotional communication features (valence, intensity, potency, and activation) and decision making options (no problem, personal problem, and significant issue) were the within-subjects factors. The rejection level for all analyses was set at $p = .05$.

2.2 Participants

Twenty-two adult volunteers participated in this study. There were 10 men and 12 women with a mean age of 34.0 years ($SD = 10.1$ years, range 23–54 years). Thirteen (59.1%) of the participants were native English speakers, while nine (40.9%) had learned English as a second language. The education level of 14 (63.3%) of the participants was postgraduate, six (27.3%) had completed an undergraduate qualification, and 2 (9.1%) of the participants had a non-university qualification. Lastly, five (22.7%) of the participants were professionals or students with a psychology, clinical psychology, or counseling background, while four (18.2%) had a medicine/nursing background. Five (22.7%) were professionals or students from social/human sciences, five (22.7%) had a background in engineering/natural sciences, and three (13.7%) selected 'other' as their field of training.

Participants were divided into two groups. Participants with backgrounds in either psychology or medicine were grouped together as possessing 'expertise in psychology' (41% of participants), while everyone else was grouped together as 'non-psychologists' (59% of participants). Participants were recruited from a university population and the Razi Educational Hospital of Psychiatry in Tabriz city, Azerbaijan, Iran. The participants from Iran were trained to an international standard in medicine, with some working in hospitals based in Melbourne, Australia.

2.3 Materials and Apparatus

Two stimuli were used in a repeated-measures design. They were segments obtained from two different, longer, video-clips, reduced to approximately six minutes in length. Both clips, labeled Scenario A and Scenario B, were of a man telling a story. The stories were different in each clip, as was the man. In Scenario A, the man was talking about his mental health issues (i.e., an eating problem and depression); while in Scenario B, the man was talking about how people might present if they have a mental health issue. Scenario B was selected as a control stimulus. Another reason two stimuli were used was to take into account individual differences in presentation styles among clients.

The stimuli were sourced from YouTube. Two steps were followed to select the clips from several available on YouTube: (1) the videos had to appear when the search terms; 'depression', 'I feel sad', and 'I feel depressed' were used, and (2) the videos could be evaluated for emotional content by members of the research team with sections of each video (e.g., 1:00 – 1:15) rated as containing 'low', 'moderate', and 'high' levels of emotion; an acceptable video had to contain all three levels of emotion to be included. Each video also had to be similar quality in terms of resolution and production values.

To assess which sources of information are important in therapeutic sessions, text-only, robotic voice, video-only, audio-only, and video-audio versions of each stimulus were constructed. Audio-only was used to mimic synchronous telephone counseling, while video-audio replicated video-conferencing sessions. Windows 7's Movie Maker was used to edit the clips. The audio content of each video-clip was transcribed verbatim, and used as a stimulus that simulated asynchronous email sessions. In addition to these major modes of delivery, robotic voices were produced from the transcripts of the stimuli; the objective of which was to remove the emotional content inherent in the human speaker's voice. BALABOLKA is a free Text-To-Speech (TTS) software and was used to generate a robotic voice of the transcripts. This software uses various versions of Microsoft Speech Application Programming Interface (SAPI) and *Microsoft Sam* male voice. The Michael voice character was used to articulate the transcripts of both stimuli artificially. The program's default volume, articulation rate, and pitch were used (Volume: 99, Articulation Rate: 0, and Pitch: 0).

Stimuli were delivered in controlled laboratory situations on the campuses of Monash and Federation University Australia as well as, Razi Psychiatry Hospital (Iran). Participants were presented with different versions of the stimuli via a Dell desktop computer with Intel(R) HD Graphic card (Display mode: 1920 × 1080, 32bit, 60Hz). Text files were opened with Microsoft Word 2010, audio files were played with the VLC media player, and video clips were watched via Windows Media Player.

2.4 Measures

The Emotionality in Cyberspace (EiC) is an instrument developed by the authors to evaluate a person's ability to perceive the emotional content of technology-based communication (see appendix). Based on previous research [30, 31, 12, 37], four subscales were included in this questionnaire. One of these, **intensity**, measures the perceived magnitude of emotions expressed by the speaker (e.g., the client). Emotional **valence** relates to the perceived quality (e.g., positive or negative) of the speaker's expression [30]. **Potency** measures the perceived strength of the expressed emotion [12, 37]. When nonverbal cues are available, actions (e.g., rotating the body away from the screen), gaze behaviors (e.g., staring), and facial expressions (e.g., lowered eyebrows) can signal how potent emotions are in speakers. **Activation** indicates how connected the speakers are (consciously) to their own emotions [36]. This factor helps determine if the emotional expression is influenced by the goals of the communication.

The EiC is a self-report measure of emotionality in which the participants rate communication on Visual Analog Scales (VASs) ranging from '0' (not at all) to '100' (a great deal). The overall experience of emotionality is calculated by summing the scores of all factors (for more details see appendix). Further, based on the study by Machado et al., differences in the clinicians' judgments of the emotional content were defined as an 'emotional deviation score' [23]. This definition provides us, first, with a way to calculate how much emotional content is observed, and second, the differences between experts' and non-experts' observations can be used to establish if there are any differences in the perceived level of emotionality transmitted via different technologies.

To test if the model was dependable, participants rated the variables (e.g., intensity) separately for each mode of delivery. Emotionality and emotional deviation scores were calculated afterward. Then, a test of the internal consistency was conducted on the data. Cronbach's coefficient for emotionality scores was 0.75 in Scenario A (representing a psychological problem) and 0.60 in Scenario B (which was a discussion about a mental health problem). Based on Cronbach's values, the internal consistency of emotional deviation scores were also checked to examine the goodness of fit which was rated high for Scenario A (i.e., 0.81), while it was estimated at a good level for Scenario B (i.e., 0.67).

The EiC instrument also included one closed and two open-ended questions. To test participants' **decision-making ability**, the closed question asked whether or not they believed that the person in each scenario had a mental health issue. Possible responses were 'no problem', 'personal issues but not significant' and 'significant issue(s) that need professional help'. The fixed, predetermined, open-ended questions were asked in a standardized order and the researcher did not deviate from the interview plan. The first open-ended question asked participants to describe, or name, the emotions they observed in each condition of computer-mediated communication (CMC). The second open-ended question asked participants to explain the types of informational markers (e.g., affective words, vocal or facial cues) they observed in the scenarios (See appendix for more details).

2.5 Procedure

After receiving ethics approval from the Monash University Human Research Ethics Committee (MUHREC), a flyer was posted on a bulletin board in the Psychology Department of Monash and Federation Universities and Razi Psychiatry Hospital inviting people to participate in the research. Upon arrival at the laboratory, each participant read an explanatory statement and completed a consent form. Participants were told that their task was to evaluate the emotional content expressed by clients in therapy-like scenarios. Participants experienced these scenarios as a written transcript, a robotic voice, a video-only, auditory information-only, and video-audio

The first six participants assessed Scenario A in the following order: text-only, robotic voice, video-only, audio-only, and video-audio conditions. This order of presentation reflected the level of non-verbal cues available in each mode, with text-only having none and video-audio having the most. After a 5 minute break, the participants then assessed Scenario B in the same order. The second six participants assessed Scenario B followed by Scenario A, with the order of conditions within each scenario the same as that for the first group. The remaining participants (n = 12) assessed the video-audio condition last, and the other conditions in a random order. The presentation of the video-audio stimulus as last was done to avoid the possibility of unduly impacting the other modes where less information was available. Participants using the EiC instrument rated each version of the stimulus. Participants were also asked the open questions and their responses were recorded. The process of completing the experiment took around 90 minutes per participant.

3 RESULTS

3.1 Analysis of the Emotionality in Cyberspace (EiC)

Descriptive statistics revealed that the communicated emotions were perceived differently across the two scenarios. The average scores were generally lower in Scenario B than they were Scenario A. In both scenarios, the valence of the communicated emotions was rated highest, while the lowest minimum average was recorded for potency. The research factors were also rated differently across differing modes of delivery. The video-audio mode was associated with the largest mean, whereas, the robotic-voice had the lowest. Table 1 reports the Means and Standard Deviations (SDs) for all factors, as well as, a summed score.

Table 1. Descriptive Analysis: Factors of Emotionality

Modes	Valence		Intensity		Potency		Activation		Summed Score	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Scenario A										
Text-based	68.09	18.44	65.04	17.67	59.09	19.34	61.59	18.93	253.81	63.13
Robotic voice	36.86	26.92	32.31	25.76	24.90	25.23	35.54	26.85	129.61	94.67
Video-only	59.59	24.81	59.95	22.37	51.45	26.06	56.68	21.14	227.58	84.77
Audio-only	70.31	19.40	70.59	16.52	60.04	21.37	66.27	18.00	267.21	65.53
Video-Audio	77.95	20.09	73.77	18.16	68.45	17.08	68.50	21.93	288.67	69.62
Scenario B										
Text-based	56.77	21.10	54.68	21.65	50.40	19.11	50.27	20.00	212.12	73.96
Robotic voice	39.40	26.18	39.36	22.85	35.45	25.73	35.90	22.68	150.11	92.56
Video-only	43.68	23.33	36.18	19.44	34.86	24.10	43.40	21.99	158.12	77.96
Audio-only	59.18	23.69	55.86	22.83	57.63	23.52	55.54	23.19	228.21	90.26
Video-Audio	62.00	23.73	59.77	23.48	60.59	24.55	62.09	23.58	244.45	85.83

3.1.1 Parametric Tests

A mixed-model ANOVA was conducted to examine if the ratings of emotional factors differed across modes of delivery, the scenarios (i.e., A vs B), and in relation to the background of the participants (i.e., psychotherapist vs. layperson). The order of modes had not influenced the measurement of variables, but there were some significant differences that are discussed in more details below.

3.1.1.1 Valence

Valence was the first factor ascribed for the tone or quality of the emotional communication. It was supposed that different technologies would impact participants' ability to perceive emotional valence. Raw data collected from the instrument for this factor was analyzed by means of a mixed-design model ANOVA. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity since Mauchly's test of sphericity was violated for the main effects of Scenario A [$X^2(9) = 19.63, p < .02$] and Scenario B [$X^2(9) = 19.62, p < .02$] across modes of delivery, as well as when the emotional valence variable was repeated between scenarios [$X^2(9) = 17.94, p < .04$].

With regard to Scenario A, which presented a potential psychological problem, the results of the analysis of the valence variable across the modes of delivery was significant, $F(2.60, 54.49) = 5.26, p < .001, \eta^2 = 0.42$. A Bonferroni-corrected post-hoc test was used to determine where the differences between the modes of delivery lay. As shown in Table 2, valence was scored significantly lower with the robotic voice than other modes ($p < .01$). Also, the comparison tests revealed that the valence rating in text-based communication ($M = 67.38; CI [59.03, 75.73]$) was noticeably reduced

compared to the video-audio mode ($M = 79.06$; $CI [70.19, 87.93]$, $p = .02$). The valence scores in video-only mode ($M = 59.33$; $CI [47.85, 70.81]$) were also significantly lower than the scores in video-audio communication ($M = 79.06$; $CI [70.19, 87.93]$, $p = .04$).

Table 2. Pairwise Comparisons of 'Valence' in Modes: Scenario A

Valence in the First Mode	Valence in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	31.25*	6.56	.001	10.57	51.92
	Video-only	8.05	5.38	1.00	-8.92	25.01
	Audio-only	-2.76	4.25	1.00	-16.14	10.63
	Video-audio	-11.68*	3.38	.02	-22.32	-1.05
Robotic voice	Text-based	-31.25*	6.56	.001	-51.92	-10.57
	Video-only	-23.20*	6.07	.011	-42.32	-4.08
	Audio-only	-34.00*	7.75	.003	-58.43	-9.58
	Video-audio	-42.93*	7.03	.000	-65.08	-20.77
Video-only	Text-based	-8.05	5.38	1.00	-25.01	8.92
	Robotic voice	23.20*	6.07	.011	4.08	42.32
	Audio-only	-10.81	6.46	1.00	-31.15	9.55
	Video-audio	-19.73*	6.09	.04	-38.92	-.54
Audio-only	Text-based	2.76	4.25	1.00	-10.63	16.14
	Robotic voice	34.00*	7.75	.003	9.58	58.43
	Video-only	10.81	6.46	1.00	-9.55	31.15
	Video-audio	-8.93	3.43	.17	-19.72	1.88
Video-audio	Text-based	11.68*	3.38	.02	1.05	22.32
	Robotic voice	42.93*	7.03	.000	20.77	65.08
	Video-only	19.73*	6.09	.04	.54	38.92
	Audio-only	8.93	3.45	.17	-1.88	19.72

Significant results were also evident in the emotional valence scores across modes in Scenario B, which was a talk about a psychological problem, $F(2.72, 57.07) = 4.96$, $p < .005$, $\eta^2 = 0.20$]. When Bonferroni-corrected tests were conducted to determine the nature of the differences between the modes of delivery, it was revealed that the robotic-voice was significantly different from the other modes. The robotic-voice mode ($M = 39.13$; $CI [27.02, 51.24]$) was associated with lower emotional valence scores than all other mediums of communication, except for the video-only mode ($M = 43.71$; $CI [32.90, 54.53]$). The results of the comparison tests between the modes are presented in Table 3. The mixed-design ANOVA was also run to test the perception of valence in the different contexts. Results showed that there were significant differences in scores for valence across scenarios, $F(3.31, 66.14) = 2.30$, $p < .03$, $\eta^2 = 0.13$. Scenario B ($M = 52.33$; $CI [45.73, 58.92]$) was scored significantly lower for valence than was Scenario A ($M = 62.40$; $CI [55.79, 69.03]$).

Figure 1 (next page) shows that the acoustic features of communication play an important role in the ability to perceive the emotional valence of a message and that video-based communication provided the most information regarding emotional valence. The profile plot illustrates a significant decrease in the ability to perceive emotional valence with robotic-voice compared with text-based and certain other modes. The ability to perceive the valence of emotions increased in the video-only, audio-only, and video-audio modes, respectively.

Table 3. Pairwise Comparisons of ‘Valence’ in Modes: Scenario B

Valence in the First Mode	Valence in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	16.87*	4.78	.02	1.79	31.94
	Video-only	12.28	5.75	.45	-5.85	30.40
	Audio-only	-3.63	4.88	1.00	-19.01	11.75
	Video-audio	-7.20	7.27	1.00	-30.10	15.71
Robotic voice	Text-based	-16.87*	4.78	.02	-31.94	-1.79
	Video-only	-4.59	7.25	1.00	-27.43	18.25
	Audio-only	-20.50*	6.17	.03	-39.94	-1.05
	Video-audio	-24.06	8.87	.13	-52.01	3.90
Video-only	Text-based	-12.28	5.78	.45	-30.40	5.848
	Robotic voice	4.59	7.25	1.00	-18.25	27.43
	Audio-only	-15.10	6.66	.27	-36.88	5.08
	Video-audio	-19.47	7.02	.12	-41.60	2.66
Audio-only	Text-based	3.63	4.88	1.000	-11.75	19.01
	Robotic voice	20.50*	6.17	.03	1.05	39.94
	Video-only	15.91	6.66	.27	-5.08	36.88
	Video-audio	-3.57	5.41	1.00	-20.60	13.47
Video-audio	Text-based	7.20	7.27	1.00	-15.71	30.10
	Robotic voice	24.06	8.87	.13	-3.90	52.01
	Video-only	19.47	7.02	.12	-2.66	41.60
	Audio-only	3.57	5.41	1.00	-13.47	20.60

The lines show that the video-audio mode of delivery remains higher than text- and audio-based communication. With that said, the negativity or positivity of expressed emotions is better transmitted by video and audio information than by audio-only or via text. Not surprisingly, the

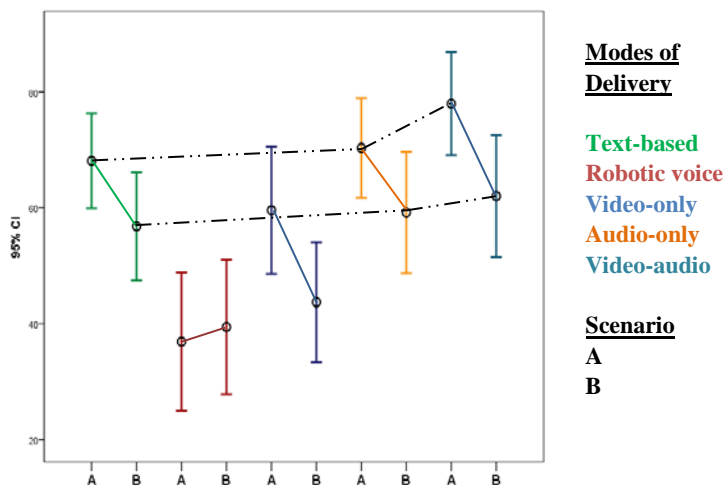


Fig. 1. Emotional Valence in Therapeutic Communication

The graph draws the mean and standard error of emotional valence scores across the modes and scenarios.

means of valence scores in the text-based, audio-only and video-audio modes of Scenario B are lower than their equivalents in Scenario A.

It was also hypothesized that participants' professional experience would be a significant factor in perceiving the level of emotional valence via CMP. Data were analyzed using a mixed-design ANOVA with a between-subjects factor of 'expertise in psychology'. There were no statistically significant differences in the valence scores, $F(2.74, 54.72) = 1.152, p < .34, \eta^2 = 0.06$. However, as shown in Figure 2, the level of expertise resulted in different nonsignificant trends between the modes. When psychology background was added to the measure, video-audio always remained the medium with the highest of perceived valence. The other modes of delivery were somewhat lower or at the same level.

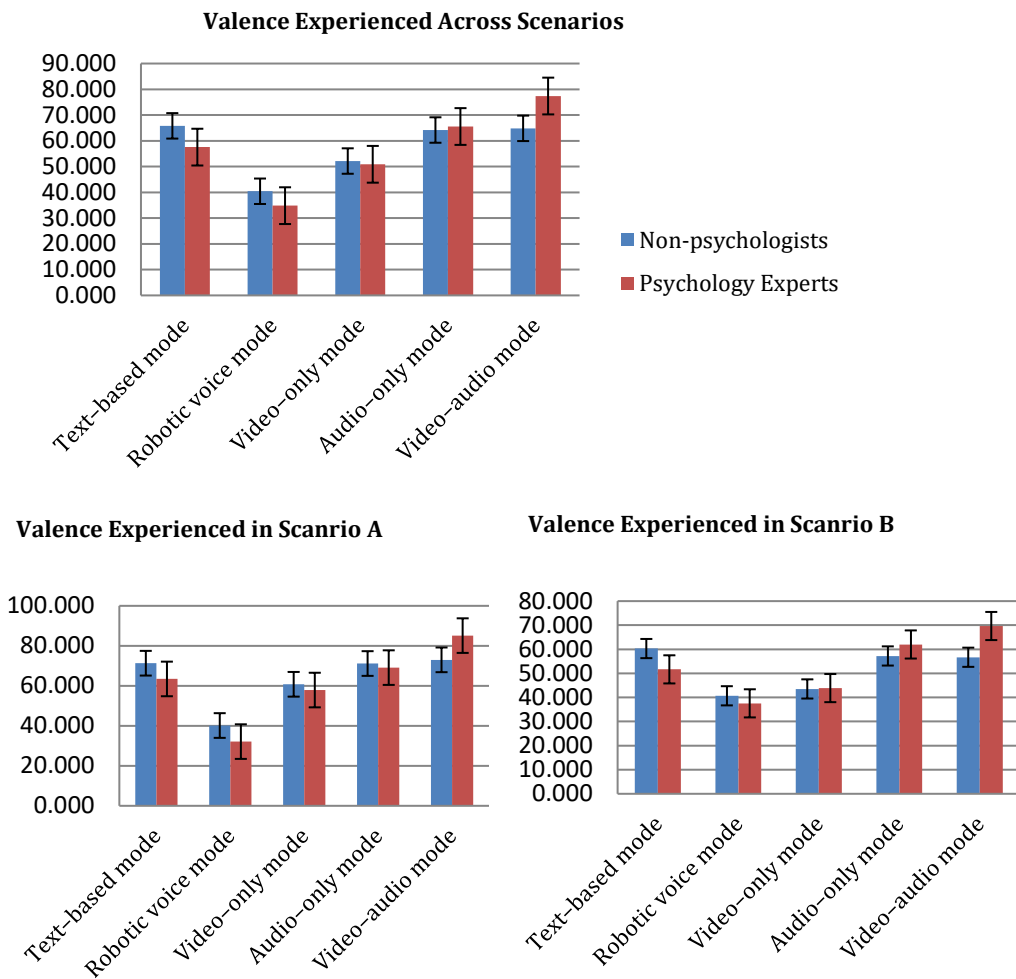


Fig. 2. Estimated Marginal Means of 'Emotional Valence': Analysis of Psychology Expertise Factor

3.1.1.2 Intensity

Intensity was the second factor which measured the perceived magnitude of the emotions expressed in communication. Similar to valence, it was assumed that intensity would be observed at different levels across technology-based contexts. A mixed-design ANOVA was used to examine the collected data on intensity. Mauchly’s test indicated that the assumption was violated in Scenario A [$X^2(9) = 24.59, p = .004$] and in Scenario B [$X^2(9) = 24.32, p < .004$]. This test was also violated for intensity scores across scenarios [$X^2(9) = 23.73, p < 0.005$]. Therefore, Greenhouse-Geisser estimates of sphericity were used in the following analysis to correct the degree of freedom.

For Scenario A, there were significant differences in perceived emotional intensity across the modes of delivery, $F(2.54, 53.41) = 5.90, p < .001, \eta^2 = 0.52$. Bonferroni-corrected post-hoc analyses were conducted to determine the nature of the differences between the modes of delivery. According to Table 4, participants observed significantly lower levels of emotional intensity in the robotic voice condition relative to all other modes of delivery ($p < 001$). The ability to perceive emotional intensity was also significantly reduced in the video-only condition ($M = 58.96; CI [48.91, 69.00]$) relative to the video-audio condition ($M = 74.91; CI [67.00, 82.81], p = .02$).

Table 4. Pairwise Comparisons of ‘Intensity’ in Modes: Scenario A

Intensity in the First Mode	Intensity in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	33.75*	5.20	.000	17.36	50.13
	Video-only	5.49	3.73	1.00	-6.28	17.25
	Audio-only	-6.12	4.08	1.00	-18.96	6.74
	Video-audio	-10.47	3.66	.09	-21.10	1.07
Robotic voice	Text-based	-33.75*	5.20	.000	-50.13	-17.36
	Video-only	-28.26*	5.56	.001	-45.78	-10.74
	Audio-only	-39.86*	6.37	.000	-59.92	-19.80
	Video-audio	-44.21*	5.56	.000	-61.74	-26.68
Video-only	Text-based	-5.49	3.73	1.00	-17.25	6.28
	Robotic voice	28.26*	5.56	.001	10.74	45.78
	Audio-only	-11.60	5.13	.35	-27.76	4.57
	Video-audio	-15.95*	4.44	.02	-29.94	-1.96
Audio-only	Text-based	6.12	4.08	1.00	-6.74	18.96
	Robotic voice	39.86*	6.37	.000	19.80	59.92
	Video-only	11.60	5.13	.35	-4.57	27.76
	Video-audio	-4.35	2.17	.59	-11.19	2.49
Video-audio	Text-based	10.47	3.66	.09	-1.07	21.10
	Robotic voice	44.21*	5.56	.000	26.68	61.74
	Video-only	15.95*	4.44	.02	1.96	29.94
	Audio-only	4.35	2.17	.59	-2.49	11.19

The main effect of mode of delivery was also examined in Scenario B (see Table 5). Although mode of delivery was significant, $F(2.83, 59.27) = 5.79, p < .002, \eta^2 = 0.22$, Bonferroni-corrected analyses showed that emotional intensity was rated as being significantly lower in the video-only mode ($M = 36.11; CI [27.10, 45.11]$) relative to the video-audio condition ($M = 60.43; CI [49.67, 71.18], p = .007$).

Table 5. Pairwise Comparisons of ‘Intensity’ in Modes: Scenario B

Intensity in the First Mode	Intensity in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	14.98	5.46	.12	-2.22	32.18
	Video-only	17.88	6.09	.08	-1.33	37.08
	Audio-only	-1.79	6.07	1.00	-20.91	17.34
	Video-audio	-6.45	7.29	1.00	-29.43	16.55
Robotic voice	Text-based	-14.98	5.46	.12	-32.18	2.22
	Video-only	2.90	7.55	1.00	-20.91	26.70
	Audio-only	-16.77	5.99	.11	-35.64	2.11
	Video-audio	-21.43	7.88	.13	-46.26	3.41
Video-only	Text-based	-17.88	6.09	.08	-37.08	1.33
	Robotic voice	-2.90	7.55	1.00	-26.70	20.91
	Audio-only	-19.67	7.18	.12	-42.30	2.97
	Video-audio	-24.33*	6.05	.007	-43.40	-5.25
Audio-only	Text-based	1.79	6.07	1.00	-17.34	20.91
	Robotic voice	16.77	5.99	.11	-2.11	35.64
	Video-only	19.67	7.18	.12	-2.97	42.30
	Video-audio	-4.66	4.22	1.00	-17.97	8.65
Video-audio	Text-based	6.45	7.29	1.00	-16.55	29.43
	Robotic voice	21.43	7.88	.13	-3.41	46.26
	Video-only	24.33*	6.05	.007	5.25	43.40
	Audio-only	4.66	4.22	1.00	-8.65	17.97

The influence of the scenario on the perception of emotional intensity was also assessed. Results of the mixed-design ANOVA showed significant differences in the ratings of emotional intensity across scenarios, $F(2.77, 55.38) = 5.83, p < .002, \eta^2 = 0.23$. As was expected, participants rated Scenario A ($M = 59.91; CI [53.32, 66.51]$) as having greater emotional intensity than Scenario B ($M = 49.06; CI [43.50, 45.61]$). As seen in Figure 3, the ability to perceive emotional intensity increased when the content was transmitted via the audio-video mode. Intensity perception in text and audio transmitted content was lower, although not significantly so. As the profile plots show,

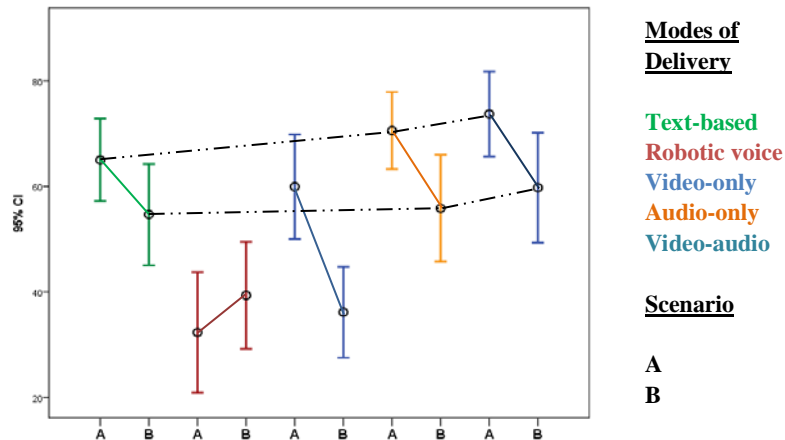


Fig. 3. Emotional Intensity in Therapeutic Communication
The graph draws the mean and standard error of emotional intensity scores across the modes and Scenarios.

the marginal means of emotional intensity has negative slopes from the text-based to the robotic voice modes. The perception of emotional intensity in the robotic voice mode declines sharply and more so for Scenario A than Scenario B. The graphs illustrate increased trends in intensity at the video-only, audio-only, and video-audio modes of Scenario A. Scenario B also showed an increase in intensity scores for the audio-only and video-audio conditions, but not the video-only communication.

Lastly, the participants' background/expertise was expected to influence the perception of emotional intensity. A mixed-model analysis showed that there were no significant influence of 'expertise in psychology', $F(2.88, 57.73) = 1.58, p < .20, \eta^2 = 0.074$. However, as shown in Figure 4, experts rated emotional intensity as being lower than non-psychologists in all modes of delivery except video-audio communicated content. Participants with a psychology background observed many more indications of intensity in this mode. This pattern is similar to their perception of emotional valence (see section 3.1.1.1).

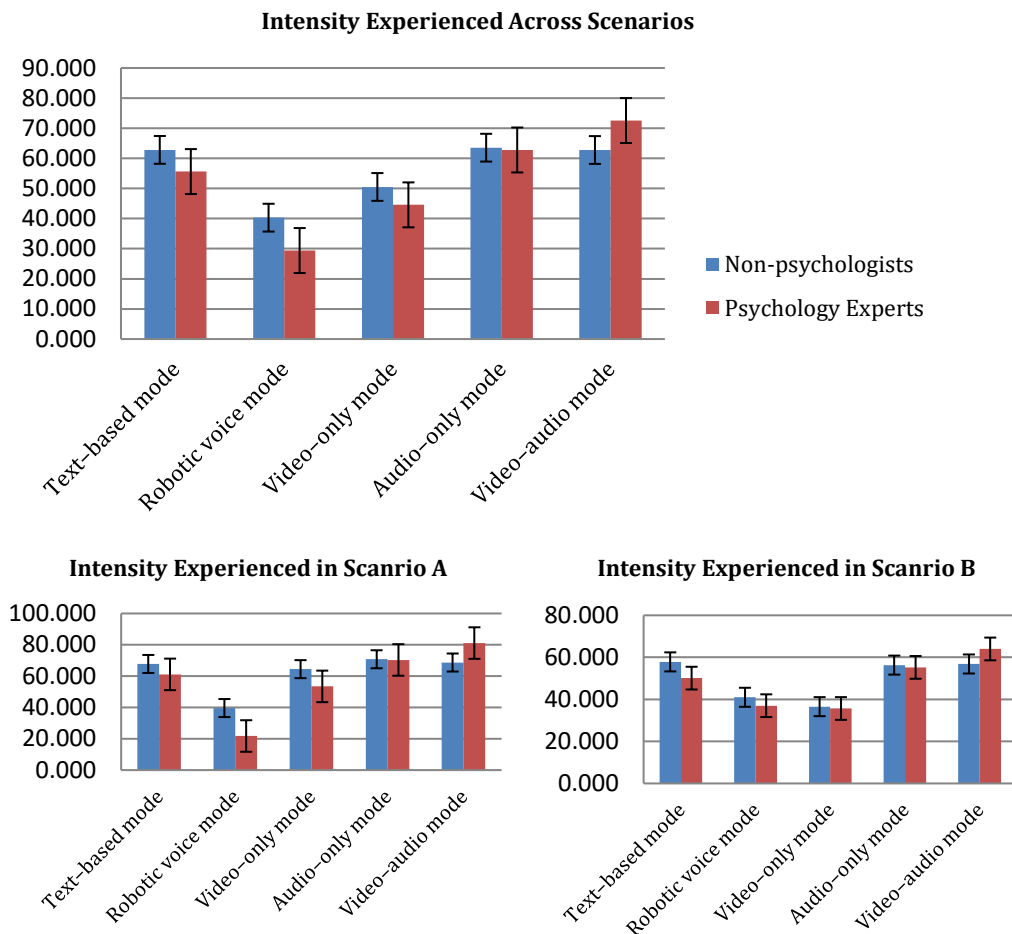


Fig. 4. Estimated Marginal Means of 'Emotional Intensity': Analysis of Psychology Expertise Factor

3.1.1.3 Potency

Potency was the third factor of the analysis and is defined as the perceived power of emotional expressions. Due to the reduction of nonverbal cues in some modes of CMC, it was assumed that perception of potency would be lower during certain modes of delivery (e.g., robotic voice). This assumption was tested using a mixed-design ANOVA. Mauchly's test indicated that the assumption of sphericity was not met in the potency scores of Scenario A [$X^2(9) = 20.57, p < .02$]. Further, the data associated with Scenario B also violated sphericity [$X^2(9) = 22.25, p < .008$]. The assumption was also violated in relation to the main effect of 'mode of delivery' [$X^2(9) = 18.40, p < .03$]. Consequently, Greenhouse-Geisser estimates of sphericity were used.

There were significant differences in the perceived emotional potency across modes of delivery in Scenario A, $F(2.71, 56.72) = 24.81, p < .001, \eta^2 = 0.54$. Table 6 presents Bonferroni-corrected post-hoc tests that showed that participants perceived more emotional potency via the video-audio mode ($M = 69.35$; CI [61.76, 76.93]) than via the text-based mode of delivery ($M = 59.33$; CI [50.85, 67.81], $p = .04$). Further, and as for other components of emotions, participants perceived the lowest level of emotional potency in the robotic voice mode ($P < .001$). Further, Bonferroni results indicated that video-audio communication ($M = 69.35$; CI [61.76, 76.93]) was scored with higher level of potency than video-only mode ($M = 50.93$; CI [38.92, 62.93], $p = .03$).

Table 6. Pairwise Comparisons of 'Potency' in Modes: Scenario A

Potency in the First Mode	Potency in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	35.58*	4.88	.000	20.22	50.94
	Video-only	8.41	4.64	.85	-6.20	23.02
	Audio-only	.29	3.73	1.00	-11.46	12.03
	Video-audio	-10.01*	3.01	.03	-19.48	-.55
Robotic voice	Text-based	-35.58*	4.88	.000	-50.94	-20.22
	Video-only	-27.17*	4.13	.000	-40.19	-14.15
	Audio-only	-35.29*	5.97	.000	-54.10	-16.48
	Video-audio	-45.59*	4.46	.000	-59.64	-31.53
Video-only	Text-based	-8.41	4.64	.85	-23.02	6.20
	Robotic voice	27.17*	4.13	.000	14.15	40.19
	Audio-only	-8.13	6.14	1.00	-27.47	11.22
	Video-audio	-18.42*	5.54	.03	-35.89	-.96
Audio-only	Text-based	-.29	3.73	1.00	-12.03	11.46
	Robotic voice	35.29*	5.97	.000	16.48	54.10
	Video-only	8.13	6.14	1.00	-11.22	27.47
	Video-audio	-10.30	3.64	.10	-21.76	1.17
Video-audio	Text-based	10.01*	3.01	.03	.55	19.48
	Robotic voice	45.59*	4.46	.000	31.53	59.64
	Video-only	18.42*	5.54	.03	.96	35.89
	Audio-only	10.30	3.64	.10	-1.17	21.76

The main effect of 'mode of delivery' was also significant for Scenario B, $F(2.79, 58.40) = 7.194, p < .001, \eta^2 = 0.26$. Different mediums were compared by using Bonferroni-corrected post-hoc tests (see Table 7). Results revealed that emotional potency was rated as being lower in the robotic voice condition ($M = 34.86$; CI [23.03, 46.69]) relative to the audio-only mode of delivery ($M = 57.66$; CI [46.76, 68.57], $p = .04$). Similarly, participants perceived a lower level of emotional potency in

the video-only mode ($M = 34.58$; $CI [23.43, 45.73]$) relative to both the audio-only ($M = 57.66$; $CI [46.76, 68.57]$, $p = .01$) and video-audio modes of delivery ($M = 60.52$; $CI [49.14, 71.90]$, $p = .004$).

Table 7. Pairwise Comparisons of 'Potency' in Modes: Scenario B

Potency in the First Mode	Potency in the Second Mode	Mean Difference	Std. Error	Sig	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	14.96	6.36	.29	-4.97	34.88
	Video-only	15.55	5.58	.11	-1.94	33.03
	Audio-only	-7.23	5.42	1.00	-24.21	9.75
	Video-audio	-10.19	6.37	1.00	-30.14	9.78
Robotic voice	Text-based	-14.96	6.36	.29	-34.88	4.97
	Video-only	.60	7.96	1.00	-24.37	25.55
	Audio-only	-22.19*	6.96	.04	-43.99	-.39
	Video-audio	-25.14	8.39	.07	-51.43	1.16
Video-only	Text-based	-15.55	5.58	.11	-33.03	1.94
	Robotic voice	-.60	7.96	1.00	-25.55	24.37
	Audio-only	-22.78*	5.99	.01	-41.53	-4.03
	Video-audio	-25.73*	6.11	.004	-44.86	-6.61
Audio-only	Text-based	7.23	5.42	1.00	-9.75	24.21
	Robotic voice	22.19*	6.96	.04	.39	43.99
	Video-only	22.78*	5.99	.01	4.03	41.53
	Video-audio	-2.96	3.42	1.00	-13.65	7.74
Video-audio	Text-based	10.19	6.37	1.00	-9.78	30.14
	Robotic voice	25.14	8.39	.07	-1.16	51.43
	Video-only	25.73*	6.11	.004	6.61	44.86
	Audio-only	2.96	3.42	1.00	-7.74	13.65

The analysis of potency scores across scenarios showed that scenario had a significant influence, $F(3.59, 71.60) = 4.15$, $p < .004$, $\eta^2 = 0.17$. Participants' identified more potency cues in Scenario A ($M = 52.49$; $CI [44.74, 60.23]$) than Scenario B ($M = 47.52$; $CI [41.10, 53.94]$).

There was considerably higher recognition of potency when a human voice and video contributed to the communication, relative to text and an automatized voice. According to profile plots of Figure 5 similar evidence of potency was found in Scenario B and Scenario A. However, the means for Scenario B were lower than Scenario A. Figure 5 also shows that the detection of emotional potency is lower for robotic voice than for text-based modes; emotional potency for the robotic voice mode shows a large decline in Scenario A. Similarly, perceptions of emotional potency in the video-only mode are lower for Scenario B compared to the text-based communication. In addition evidence of emotional potency in the audio-only and text-based modes is similar. Again the video-audio mode has greater perception of potency.

Similar to the other emotionality factors, it was assumed that participants with a background in psychology would perceive different levels of emotional potency than non-experts during CMP sessions. However, the split plot of ANOVA did not show significant difference when 'expertise in psychology' was employed in the measurement, $F(2.70, 53.54) = 0.70$, $p < .60$, $\eta^2 = 0.34$. However, as can be seen in the profile plots in Figure 6, psychologists noted fewer cues for potency except when the content was transmitted via video-audio communication.

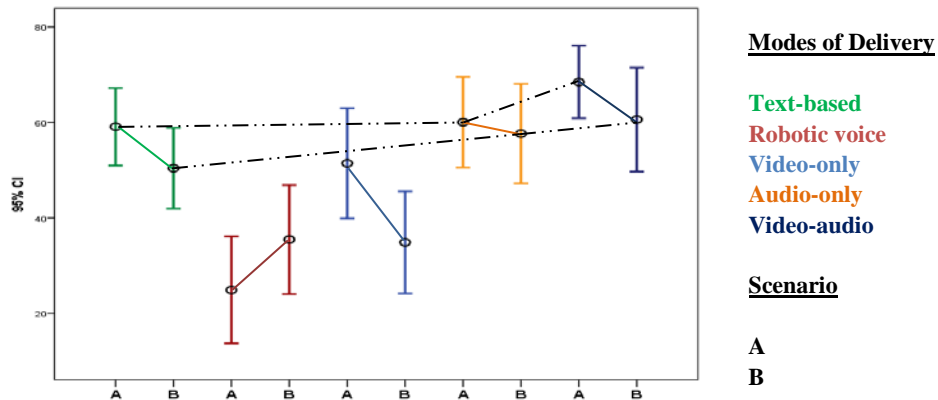
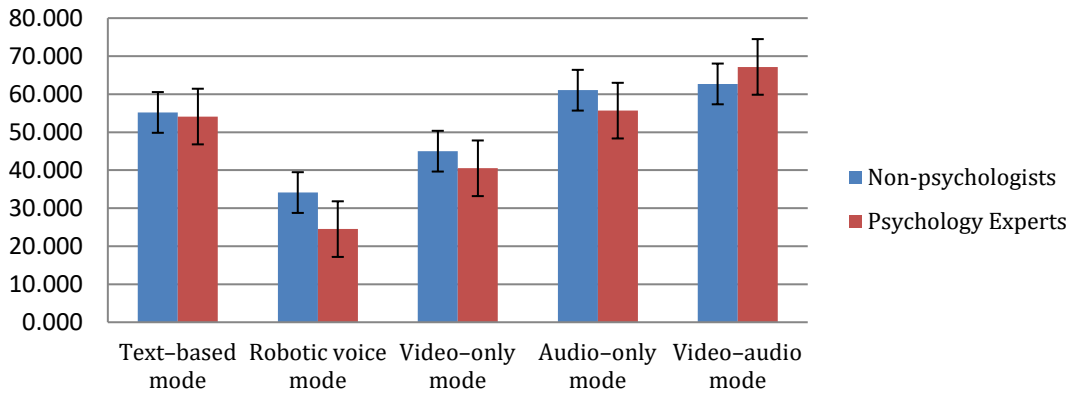
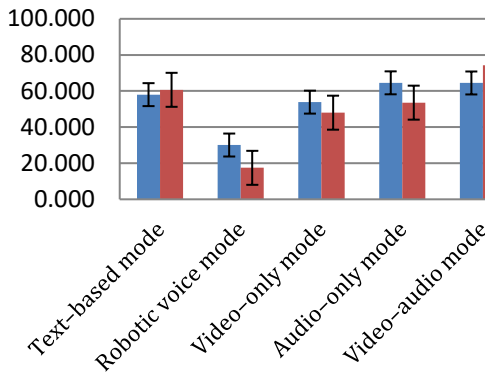


Fig. 5. Emotional Potency in Therapeutic Communication
The graphs include the mean and standard error of emotional potency scores across the modes and scenarios.

Potency Experienced Across Scenarios



Potency Experienced in Scenario A



Potency Experienced in Scenario B

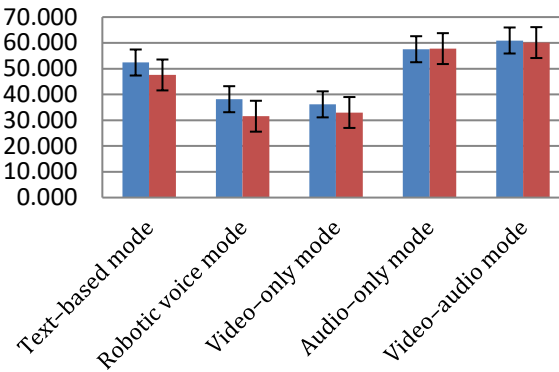


Fig. 6. Estimated Marginal Means of 'Emotional Potency': Analysis of Psychology Expertise Factor

3.1.1.4 Activation

Activation was the last factor for assessing emotionality by scoring the amount of emotional (action) tendency or affective motivation in communication. To test the assumption that the perception of emotional activation varied across different technologies, mixed-design ANOVAs were employed. The results showed that data associated with Scenario A violated the assumption of sphericity [$X^2(9) = 25.80, p < .002$]. This test was also true for Scenario B, [$X^2(9) = 17.80, p < .03$]. The assumption was also violated when activation scores repeated across the scenarios [$X^2(9) = 18.60, p < .03$]. As such, Greenhouse-Geisser estimates of sphericity were employed.

A repeated-measures ANOVA revealed that participants' perceptions of emotional activation were significantly different across modes of delivery in Scenario A, $F(2.33, 55.58) = 12.83, p < .001, \eta^2 = 0.39$. To determine the difference between the modes of delivery, Bonferroni-corrected post-hoc tests were employed. The tests showed that in this scenario, the robotic-voice mode was significantly different from all other modes of delivery (see Table 8). This suggests that this mode of delivery impedes participants' ability to perceive emotional activation.

Table 8. Pairwise Comparisons of 'Activation' in Modes: Scenario A

Activation in the First Mode	Activation in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	27.047*	6.896	.008	5.301	48.793
	Video-only	5.350	5.592	1.000	-12.282	22.983
	Audio-only	-3.654	3.999	1.000	-16.265	8.957
	Video-audio	-8.325	2.981	.112	-17.724	1.075
Robotic voice	Text-based	-27.047*	6.896	.008	-48.793	-5.301
	Video-only	-21.697*	4.968	.003	-37.361	-6.032
	Audio-only	-30.701*	6.991	.003	-52.745	-8.656
	Video-audio	-35.372*	6.723	.000	-56.572	-14.172
Video-only	Text-based	-5.350	5.592	1.000	-22.983	12.282
	Robotic voice	21.697*	4.968	.003	6.032	37.361
	Audio-only	-9.004	5.124	.942	-25.162	7.153
	Video-audio	-13.675	5.819	.292	-32.026	4.676
Audio-only	Text-based	3.654	3.999	1.000	-8.957	16.265
	Robotic voice	30.701*	6.991	.003	8.656	52.745
	Video-only	9.004	5.124	.942	-7.153	25.162
	Video-audio	-4.671	3.691	1.000	-16.311	6.969
Video-audio	Text-based	8.325	2.981	.112	-1.075	17.724
	Robotic voice	35.372*	6.723	.000	14.172	56.572
	Video-only	13.675	5.819	.292	-4.676	32.026
	Audio-only	4.671	3.691	1.000	-6.969	16.311

The mixed-design ANOVA test was also repeated for Scenario B. The results showed that the main effect of mode of delivery was significant, $F(2.79, 55.65) = 5.21, p < .001, \eta^2 = 0.21$. Therefore, Bonferroni-corrected post-hoc tests were run to identify the nature of this difference. According to Table 9, there were significant contrasts between the following modes: 1) the robotic-voice ($M = 35.43$; CI [24.99, 45.87]) and video-audio ($M = 61.91$; CI [50.99, 72.83]) and 2) the video-only ($M = 43.21$; CI [33.03, 53.40]) and video-audio modes of delivery ($M = 61.91$; CI [50.99, 72.83]). The findings suggest that, in the absence of verbal features, participants scored a reduced level of perception for the emotional activation of communication.

Table 9 Pairwise Comparisons of ‘Activation’ in Modes: Scenario B

Activation in the First Mode	Activation in the Second Mode	Mean Difference	Std. Error	Sig.	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Text-based	Robotic voice	14.28	5.59	.19	-3.33	31.88
	Video-only	6.50	5.36	1.00	-10.39	23.39
	Audio-only	-5.92	6.52	1.00	-26.47	14.64
	Video-audio	-12.21	6.67	.83	-33.24	8.83
Robotic voice	Text-based	-14.28	5.59	.19	-31.88	3.33
	Video-only	-7.79	7.78	1.00	-32.15	16.59
	Audio-only	-20.19	7.01	.09	-42.28	1.90
	Video-audio	-26.48*	8.16	.04	-52.19	-.78
Video-only	Text-based	-6.50	5.36	1.00	-23.37	10.39
	Robotic voice	7.79	7.73	1.00	-16.59	32.15
	Audio-only	-12.41	6.30	.63	-32.28	7.46
	Video-audio	-18.70*	5.72	.04	-36.72	-.69
Audio-only	Text-based	5.92	6.52	1.00	-14.64	26.47
	Robotic voice	20.19	7.01	.09	-1.90	42.28
	Video-only	12.41	6.30	.63	-7.46	32.28
	Video-audio	-6.30	4.18	1.00	-19.46	6.89
Video-audio	Text-based	12.21	6.67	.83	-8.83	33.24
	Robotic voice	26.48*	8.16	.04	.78	52.19
	Video-only	18.70*	5.72	.04	.69	36.72
	Audio-only	6.30	4.18	1.00	-6.88	19.46

The main effect of “scenario” was significant, $F(2.72, 54.39) = 13.17, p < .001, \eta^2 = 0.40$. When scenarios were compared, the results show that participants identified activation more frequently in Scenario A ($M = 57.49, CI [50.93, 64.04]$) than in Scenario B ($M = 49.18, CI [43.30, 55.06]$).

Similar to other factors, the lines in Figure 7 illustrate slight changes in the perception of emotional activation across text, audio, and video-audio modes. However, in comparison to the text,

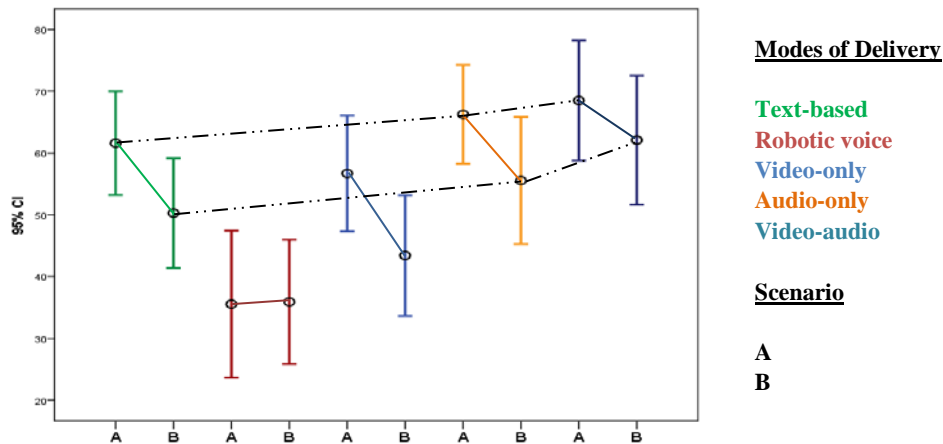


Fig. 7. Emotional Activation in Therapeutic Communication
The graph draws the mean and standard error of emotional activation scores across the modes and scenarios.

the audio-only enhanced observation of emotional activation more frequently while the video-audio mode still enabled the greater recognition of activation. For perception of emotional activation, communication that included human audible and visual components was more favorable. Also, activation was perceived more often in the text-based than robotic voice modes. The robotic voice also attracted considerably lower scores than the video-only mode. Thus robotic voice communication was not documented as a good medium for perceiving emotional activation. Ultimately, according to this figure, the pattern of perceiving emotional activation looks similar to the other factors in emotional communication

Finally, the data for emotional activation was explored with respect to the participants' psychotherapy background. Based on the mixed-model ANOVA (see Figure 8), scores were not significantly different across scenarios when the between-subjects factor of 'expertise in psychology' was introduced, $F(2.72, 54.39) = .75, p = .53, \eta^2 = 0.04$. However, profile plots in Figure 8 show some different nonsignificant trends among participants. Those with a psychology background rated emotional activation as being less apparent than did non-psychologists in all modes except the video-audio mode, where experts recorded a higher score for emotional activation.

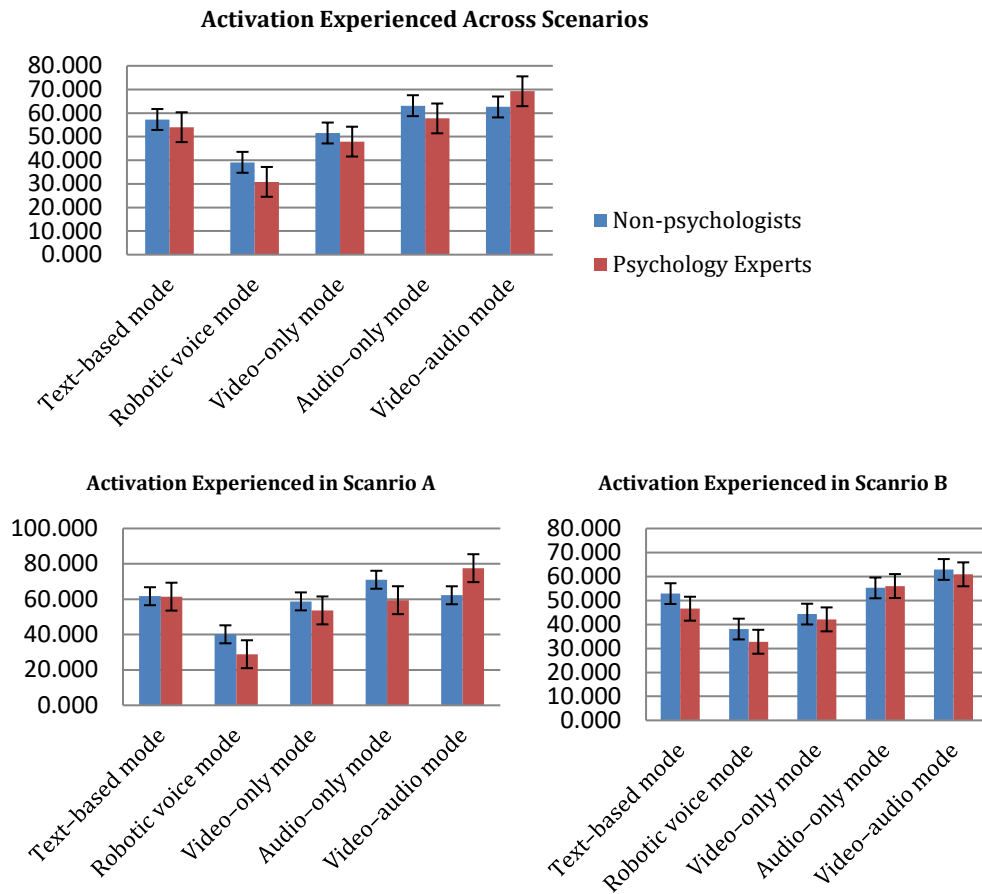


Fig. 8. Estimated Marginal Means of 'Emotional Activation': Analysis of Psychology Expertise Factor

3.2 Self-Reports of the Participants about the Emotionality in Cyberspace (EiC)

In addition to quantitative data, participants were interviewed and asked to contribute their perspectives on the emotional content of the communication in each scenario. Key interview findings are described below.

When participants were asked to name or describe the emotions they had perceived in the online communication tasks, negative emotions were most often perceived in Scenario A and across all modes of delivery. For example, sadness, shame, hopelessness, upset, uncomfortable, confusion, frustration, uncertainty, anxiety, aloofness, discouragement, dismay, low self-esteem, tiredness, depression, apathetic, ambivalence, testiness or petulance were used to describe the emotions presented in Scenario A. The second scenario, which included a talk about a mental health problem, was mostly nominated as having positive emotions. Confidence, curiosity, interest, pride, enthusiasm, hope, determination, and felicity are examples of the positive emotions used when describing this scenario; while mention was also made of anxiety, fear, and embarrassment by some of the participants (8%).

When participants were asked to describe how these perceived emotions had been conveyed in the scenarios and across modes, the majority of participants had responded to affective language. Direct emotional words (e.g., bad moods, hopeless, scared) and indirect affective phrases or sentences (e.g., 'I can't control myself' and 'no one takes me seriously') were referred to as the cues in the text-based condition. One participant stated that the text-based mode in Scenario A displayed intense depression due to sharp and straightforward psychological/clinical words or expressions such as 'I do not have depression', 'I am really pissed off' or 'I do not need to be this guy'.

Three other themes were highlighted as important by participants with regard to text-based communication. First, sociocentric phrases that regard one's own social orientation in a group were emphasized as the first textual cues that contributed to the ability to perceive the potency and activation of the emotional states of the speaker. According to one participant, the client in Scenario B using sociocentric expressions such as, 'be true to yourself' or 'you know' was interpreted as having self-confidence, determination, and being able to provide intellectual advice. The second textual cues mentioned were made of egocentric / monotonic conversation that regards thinking only of oneself and one's feelings or desires. One participant, a trained counselor, believed that her diagnosis of depression in the client of text-based condition of Scenario A was associated with increased use of the first-person, singular pronouns (e.g., 'I', 'me,' and 'myself'). In her opinion, this egocentric cue signaled activated and intense mental problems in the speaker. Rationalizing past events, feelings or thoughts was deemed the third way to demonstrate emotional meaning, but not emotional feeling. For example, if the speaker used phrases such as 'I have the right to live' or 'I was in a dark place' the participant felt that the speaker was talking to an audience empathetically, or wanted to provide helpful advice by means of his/her past experiences.

Emotional words/phrases were the only cues available to understand what the robotized speaker was feeling, but these only provided a generalization of emotion (e.g., anxiety). Although emotional words/phrases were highlighted as the important cue, participants argued that the artificial articulation was rigid and impassive for emotional arousal. In other words, the lack of vocal inflections was the main hurdle in the ability to identify emotions. There was a participant, for example, who believed that she could understand all that was said but the lack of tone, pitch, and pauses made the robotized communication very unemotional. Similarly, other participants argued that the robotic mode of delivery embodied little emotional information, and thus provided no information regarding the intensity, activation, and arousal of any emotions expressed in the communication. Further, all participants stressed a lack of connection with the communicator when presented with the robotic voice. Consequently, the synthesized voice did not evoke much empathy for, or much of an emotional alliance with, the client.

Participants' statements suggested that a real human voice rather than a robotic voice made perceiving emotions easier. Vocal cues in the audio-only mode (such as a monotone voice, long pauses, expressing phrases at a different pace) were thought to be cues of greater emotionality. Gaps between particular emotional phrases also suggested contemplation and resignation on the part of the speaker. The tone or intonation pattern in the audio-only mode was also thought to be a cue that carried a large proportion of emotional information. One participant, for example, believed that although the articulation of words or sentences in Scenario B did not reflect an emotional state, they did suggest nervousness on the basis of speed (i.e., the words were run quickly by the speaker). In contrast, a slow rate of speech was deemed to be an important indicator that the speaker was experiencing a flattened affect. This was identified as a symptom of depression in Scenario A. Among the tonal features of the voice, the pitch or rate-of-speech were considered significant for psychotherapy purposes. For example, while upset or distressed moments accompanied a low-pitch voice in Scenario A, tearfulness and a catching in the throat were evident when the pitch was raised.

Video-based delivery was deemed to be a more reliable medium in the perception of emotional content. Participants noted that the integrated information available as part of video-based communication (i.e., affective language, voice, and observable behavioral factors) aided them to better perceive emotional content. For example, the potency of distress or upset apparent in Scenario A was identified by participants through affective words and occasional gestures such as, scratching the head, touching the face or mouth, or moving the chin from side-to-side. Facial expressions were also considered important in determining emotions. For example, negative feelings were exemplified by frowning and crying. Eye contact provided a further cue for emotional perception. The speaker in Scenario A, for example, did not maintain gaze towards the screen demonstrating less of an intense emotional connection with the audience relative to the speaker in Scenario B, this was perceived as a symptom of depression. Finally, positioning oneself at the edge of the screen or looking directly into the camera was interpreted as potency, valence of determination, and confidence; this behavior also represented decreased critical emotional activation.

Although CMP environments that provided multiple sources of information were regarded positively, some concerns were raised. Participants noted that, in general, it was much harder to perceive emotionality via virtual transmissions, relative to their memories of face-to-face communication. They felt that there was always something they could not discern from the screen, or the flatness of the screen did not present the speaker as life-like. Certain body movements, such as leg tremors, chest movement (e.g., the rise and fall of the chest associated with breathing) or small movements of the fingers, were unobservable. Some participants mentioned that a depressed person tends to sigh frequently, while an anxious person will often tap their foot or rub their sleeves, and these movements could not be observed on the screen or in other environments. Based on this limitation, participants agreed that they preferred face-to-face counseling sessions over computer-mediated modes of psychotherapy.

3.3 Diagnosis of Psychological Problems and Emotionality of Communication

The final line of inquiry in this study was to test whether mode of delivery influenced participants' judgment about the mental health status of the communicator. It was assumed that any diagnoses made by participants would be affected by the technology used. Information regarding a clients' behavior is crucial for diagnosis, so these diagnostic skills may be challenged with CMP. In the EiC measurement, to examine this assumption, participants were asked whether the therapeutic content communicated through various modes of delivery (i.e., the robotic voice, text-based, audio-only, and video-audio modes) provided evidence of either 'a significant mental problem that needed professional help' or 'a personal issue that didn't require treatment' or that a 'mental problem was not recognized in the speaker'. These levels were grouped under the variable 'diagnosis of

psychological problem' for the purpose of the analysis. The results of these descriptive statistics are described as follows:

Scenario A

- The results showed that Eleven (50%) participants believed that the speaker has a 'significant mental health issue' in the text-based communication mode. Of the 11, seven out of 9 participants had a background in psychology, while four out of 13 had no psychological training.
- Nine (40.9%) participants reported a 'significant mental issue' in the robotic voice condition. Six out of 9 of these participants had a background in psychology, while three out of 13 of participants did not.
- Six (27.3%) participants identified the speaker as having a significant mental health concern in the video-only mode. Three out of 9 of these people were experts in psychology, while three out of 13 participants had no training in psychology.
- When vocal cues were present (i.e., audio-only communication), 14 (63.6%) participants reported that the speaker had a significant mental health issue. All participants with a background in psychology (n = 9) suggested that the speaker had a significant mental health issue when using this medium.
- When the video-audio mode was used, the number of participants who suggested that the speaker had a significant mental health concern remained the same (14 participants; 63.6%). All participants with a background in psychology (n = 9) and 5 participants who had no expertise in psychology reported a significant issue.

Scenario B

- 50% percent of the participants (n = 11) selected 'significant mental issues that need professional help' with the text-based mode which is similar to Scenario A. Six out of 9 of these participants had a psychology background, while five out of 13 had no psychology training.
- The 'significant mental issue' in robotic voice was identified by 8 (36.4%) participants. Two out of 13 of these participants had no psychology training, while six out of 9 were experts in psychology.
- Descriptive statistics for the video-only mode showed that a reduced number of participants reported a mental health concern, relative to Scenario A. Two (9.1%) participants recognized a 'significant mental issue' with this medium. One participant was from a group of 9 who had a psychology background, while the other was a person without psychology training from a group of 13.
- The audio-only mode was marked with 'significant mental issues' by 7 (31.8%) participants. This medium was identified by three out of 13 non-experts in psychology, while four out of 9 participants had psychology training.
- Six (27%) participants identified the speaker of the video-audio mode with a 'significant mental issue'. Similar to the audio-only medium, the current mode had the same number of non-experts (i.e., three out of 13) who arrived at this judgment, but participants with psychology background reduced to three out of 9.

These decision-making data were then compared to examine the influence of alternate technologies on the 'diagnosis of a psychological problem'; the data, however, were not analysed across scenarios because they communicated different and independent contexts. A Friedman test showed no significant differences in the data collected for the main modes of delivery in Scenario A ($p > .2$), suggesting that participants made similar decisions on the severity of mental problems

via text-only, audio-only, and video-audio communication. However, the Friedman test showed significant differences in the number of participants who selected 'diagnosis of a psychological problem' in the robotic-voice and audio-only modes, $X^2(22) = -2.31$, $p = .02$, and between the video-only and video-audio modes, $X^2(22) = -2.49$, $p = .01$. There were significantly fewer participants who identified a mental health issue in the robotic-voice and video-only modes, relative to the audio-only and video-audio modes, respectively.

The Friedman test for Scenario B showed that participants made significantly different decisions regarding the speaker's mental health issues in text-based and audio-only communication, $X^2(22) = -2.18$, $p = .03$. Significant results also revealed that more participants (50%) diagnosed the speaker as having a significant mental problem in the text-based communication compared with the video-audio mode of delivery, $X^2(22) = -2.49$, $p = .01$. This diagnosis was less frequent via the audio-only and video-audio mediums of delivery (31% and 27%, respectively).

As such, these findings contribute evidence suggesting that diagnostic accuracy can be improved if verbal and visual cues are added to CMP.

3.3.1 The Impact of Emotionality on Decision-Making

Lastly, a regression analysis was conducted to assess whether or not the emotionality of therapeutic communication influenced diagnostic decision-making across modes of delivery. A dummy coding system was used to compare the dependent variable 'a significant mental issue' with the predictor variable 'emotional deviation scores', ΔE . The emotional deviation score formulated the difference between those with and those without a background in psychology on a certain aspect of emotional perception to derive emotionality of the communication (see Section 2.3; [70]). In results obtained for Scenario A, the ΔE showed a significant proportion of variance in the diagnosis data via the text-based mode, $F(1, 21) = 6.59$, $p < .02$, $R^2 = .25$, $R^2_{Adjusted} = .21$. Thus, reduced emotionality may predict a significant decline in the diagnostic ability of psychotherapists via this mode, $\beta = -.50$, $t(21) = -2.57$, $p < .02$. The ΔE scores explained a significant variation in diagnosis data via the audio-only mode, $F(1, 21) = 8.33$, $p < .009$, $R^2 = .30$, $R^2_{Adjusted} = .25$. A significant decline in diagnostic ability was also evident in this mode relative to reduction in the scores of ΔE , $\beta = -.54$, $t(21) = -2.89$, $p < .009$. A significant proportion of variance was also observed for diagnosis data with ΔE scores of the video-audio mode, $F(1, 21) = 10.54$, $p < .004$, $R^2 = .35$, $R^2_{Adjusted} = .32$. Thus, a significant negative relationship between emotionality, ΔE , and diagnostic ability was also reported for the video-audio mode, $\beta = -.59$, $t(21) = -3.25$, $p < .004$.

For Scenario B, participants' diagnosis of a 'significant mental issue' was not significantly dependent on the ΔE scores across any mode. This was probably due to the fact that Scenario B contained little emotional content.

Finally, the reduction in the ability to perceive emotional content in several telecommunication modes can negatively impact decision-making. Specifically, the results of this study suggest that text-based and robotic-voice communication impair clinicians' ability to accurately diagnose clients.

4 DISCUSSION

The central aim of the current study has been to examine how emotionality in therapeutic communication is perceived via technology. Due to the different levels of verbal and nonverbal cues in computer-mediated communication (CMC; [38]), it was presumed that clinicians would perceive emotional content differently in the text-based, audio-only, and video-based conditions. Overall, the study suggests that different computer-mediated environments can lead to differences in perceiving communicated emotions.

According to the results, the perception of emotional intensity (i.e., the perceived magnitude of the expressed emotions) varied across the modes of delivery. The emotion content of the information was less available to participants in the robotic voice condition. Further, and consistent with the findings of Machado et al. [23], the current results show that the intensity of emotions was rated lower in the text-based conditions than it was in other conditions. Similar to Chang Lee, Kwon, Kim, and Kim [39], this finding suggests that therapeutic messages are less emotionally intense when presented via a text-based, asynchronous communication.

In other modes of delivery, and as is evident in the literature [40], integration of a human voice with facial expressions in computer-mediated psychotherapy (CMP) made no significant difference to the ability to perceive emotional content, although there was a slight increase in the degree of rated emotional intensity. For example, a decreasing correlation was found for the ability to perceive emotional intensity in the audio-only and text-based modes of delivery. The audio-only mode was associated with stronger emotional information. When compared with the video-audio mode, audio-only communication stimulated a lower level of perceived emotional intensity. It appears as though the intensity of a client's emotional state was better perceived when determiners had access to verbal communication, body language, and facial expression. Thus, these findings suggest that while audio is better than text, video is better than audio for perceiving emotional intensity, but not other emotional elements, of communication.

The present study also found that synthetic emotional expressions impeded psychotherapeutic communication between the client and the therapist. The robotic voice mode prevented participants from accessing information on the valence of emotions. Further, and consistent with our assumption, therapeutic communication via text showed the lowest level of perception of emotional valence. Therefore, affective language including emotional words, as would occur in text, was not adequate for the prediction of emotional valence. Nonetheless, participants were able to distinguish types of (positive or negative) emotions in the text-based mode and in the different scenarios: negative emotions for Scenario A and positive emotions for Scenario B. This finding is consistent with previous studies by Hancock, Landrigan, and Silver [41] who showed that types of emotion are distinguished properly in text-based communication.

The observation of valence was different in other modes of delivery. Emotional valence was scored higher in the audio-based communication than in text messages. This finding is consistent with research by Nass, Foehr, Brave, and Somoza [42], who showed that valence conveyed by the voice was adequately perceived by clinicians. Nass et al. (2001) also concluded that the human voice had a greater influence on clinicians' perceptions of emotional valence than a synthesized voice. In this study, valence was distorted by the robotic voice and it was not possible to distinguish between negative and positive emotional cues without the prosody of the human voice. The current study also shows that the most prominent presentation of emotional valence occurs with video-based communication, suggesting that embodied information including bodily movements contribute considerably to the assessment of this feature.

Emotional activation (action tendency) and potency (power of emotional expressions) were also perceived differently across modes of delivery. In contrast to intensity and valence, the potency and activation of emotions had generally lower scorings through text-, audio-, and video-based communication. While less activation and potency was identified in text messaging, this was particularly so with the robotic voice. In contrast, these emotional cues were more evident to listeners when a human voice was used to provide acoustic information, and this improvement was increased when information was provided audio-visually. Participants suggested that positive and negative potency content was easier to identify when both speech and facial expression could be assessed. For example, negative content was represented by slower body movements, while positive content was identified via certain vocal expressions such as normal pitch. Thus, video information enhances perception of the strength and activation of clients' concerns, intentions, and feelings,

whereas an artificial voice does not value how strong a cyber-client feels at the time of interaction. Artificial voice also would not demonstrate if cyber-clients have dispositions to take some actions in relation to their feelings.

The results also show that participants with a background in psychology were more perceptive than non-psychologists in identifying emotions in technology-based communication. The experts in psychology interacted carefully with the intensity of expressed emotions in the therapy-like scenario. This is consistent with the findings of Machado et al. (1999), who found that therapists could identify emotions or feelings of clients more accurately than non-therapists in videotaped conditions [23]. However, in that study the ability to identify emotions was less reliant on verbal cues among psychotherapists than among non-therapists. Findings from interviews in this study showed that participants with psychology backgrounds relied more on verbal and behavioral cues to perceive emotions than did non-experts.

Another finding that provides practical guidance on assessing the strength of emotions is the affective wording used by clients. For instance, participants of the current study explained in their interviews how affective words contributed to understanding the intensity and valence of emotional content in a therapeutic message. Direct emotional words (e.g., happy, sad, sorry) were stressed as they can denote positive or negative emotional states. Quantifying words (e.g., much, many, a lot) were used to assess the intensity of the emotions expressed. Indirect emotional phrases, on the other hand, such as 'a dark road' or 'a tunnel without light' were proposed by participants as metaphors indicative of emotional valence, such as an aspect of depression [43].

According to the participants, other examples from the client's language (e.g., the frequent use of the first person pronoun 'I') can express possible causes of emotional arousal, such as sadness in a message. These can represent certain psychological states or even expose mental health issues such as depression. No such interpretation was attached to second-person pronouns (e.g., 'you'), which were frequently used by the clients (the speaker in Scenario B). These findings are consistent with other recent research. In their studies, Brockmeyer et al. [44] and Zimmermann et al. [45] demonstrated that the frequent use of first-person singular pronouns was associated with heightened depressive symptoms. The use of first-person pronouns was interpreted as heightened interpersonal distress, while first-person plural pronouns (e.g., 'we') was associated with fewer depressive symptoms. Notably, the current study adds to this research by suggesting that this relationship between first-person singular pronouns and depressive symptoms exists across different modes of communication.

Participants also explained the role of prosody in assessing a client's situation. The intensity of sound (i.e., the pitch and volume) and the speed of articulation were considered to be the main indicators for distinguishing intensity, activation, and potency of clients' emotional states. This is consistent with studies of acoustics features in emotional expressions [46, 47]. For example, there are influential associations between vocal cues (e.g., rate of speech, loudness) and emotions such as fear anxiety, sadness, and depression [48]. Participants in the present study explained that slower and softer speech was indicative of depression, but a normal rate of speech and a normal tone indicated that the speaker had positive thoughts. This suggests that vocal cues provided participants with emotional information that the text-only mode did not. This interpretation also provides a reasonable explanation as to why using an artificial voice may not be useful for psychotherapeutic communication, especially given the quantitative findings presented above.

Other behavioral cues provide information about a clinician's emotional engagement with cyber-clients. For example, visual contact is important for assessing emotionality [49]. In this study, gestures played an important role in the participants' assessments (e.g., eye movements, frowning, crying, and closeness to the screen). Further, head and hand movements (e.g., rocking the head from side-to-side and up-and-down, touching one's body and/or head, scratching, pulling hair, and tapping) were also deemed important. The observation of behavioral features, however, can be

hindered via technology. For instance, some of the body movements (e.g., foot tapping), which are important for diagnoses, were not available in the video-based modes of delivery. Participants also reported that they could not distinguish eye movements on the screen as well as in face-to-face sessions. This can mean a lower level of empathy, connectedness, or emotional relevancy with a client via videoconferencing [50]. In sum, the absence of these important behavioral elements may reduce the perception of emotionality in communication and impair the CMP process.

To focus further on the construct validity of emotionality in CMP session, the secondary aim of the study was to examine if heuristic diagnoses can be performed through technology-based environments. Heuristic decision-making is described in the literature as a process, which takes place with little information [51, 52]. A heuristic approach can provide less examination error and more transparency in diagnosis [27]. In this experiment it was assumed that the heuristic diagnostic ability of psychotherapists would be affected by CMP due to reduced cues associated with emotional and behavioral symptoms. It was found that participants changed their therapeutic decisions across the different delivery modes. In partial support of the hypotheses, the intensity of emotions was not recognized adequately via the robotic voice mode. This suggests that using non-human voices in audio-based psychotherapy should not be recommended for mental health purposes.

However, contrary to the hypotheses, the other CMP conditions (i.e., modes other than robotic voice) did not contribute to differing abilities to diagnose. Nonetheless, there were main points of difference between them. The statistics indicated that text-based communication, which was perceived as having a lower level of emotionality, had relatively few participants who agreed on the severity of the mental problem. This finding supports that of a previous study by Bambling et al. (2008), who showed impairment in text-based diagnosis due to the reduced intensity of emotional communication [29]. Unlike text, when video-audio information was presented the diagnostic ability increased. The diagnosis of significant mental health issues in video-based mode was more accurate than in audio-only communication. This suggests that videoconferencing may be the best substitute for face-to-face therapy. In previous empirical studies, this mode has been demonstrated as providing more accurate information for assessment of mental disorders such as dementia [53].

Finally, the results showed that participants with knowledge in psychology rated the mental health issues in the robotic voice and text-based modes as being more significant than those without psychology experience. This suggests that education and training in psychotherapy can help overcome some of the limitations associated with certain computer-mediated environments. Additionally, the decision-making of the trained participants showed a relationship with their understanding of emotionality indicating that trained cybertherapists can successfully decode CMP communication and perform heuristic diagnoses.

4.1 Major Limitations

This research contains several limitations, a number of which need to be acknowledged. In addition to the limited number of professional participants, in the current study there was no reciprocal communication between counseling pairs [54]. As such, the complexity of relationships in cyber-clinical contexts was ignored. Additionally, conditions of delivery ought to be adjusted to include the perception of emotionality with real clients. The stimuli used in the current study may not represent the reality for cyber-clients. The strong ethical, technical, and financial barriers against recruiting real patients forced the researcher to employ role-play and simulated clients. Besides those already mentioned, current measure of EiC may not yet be used as a generalized instrument and should be inspected for the strength of validity and reliability.

4.2 Future steps

Although the current study offers exploratory insights, further research is needed to support the current approach to addressing the problem of perceiving emotionality in CMP communication. The

first need regards probable biases in the current inferential statistics. Indeed, a larger sample of participants, particularly mental health providers, could help avoid some of the pitfalls identified here. A further step should be the examination of emotionality in different real situations – individually or in group therapy. To enhance the validity of the findings, it is also necessary to consider individual psychological problems (e.g., bulimia, PTSD) when communicated via technology. The findings presented here could be enhanced if emotional coding structures are used in future studies [1]. Emotion coding techniques assist psychotherapists to label the feelings, concerns, and intentions of clients and provide efficacious treatments [55]. Finally, the computer-mediated conditions discussed in this study did not exhaust all possible CMP modes. In the future, it would be important to examine emotionality through other mediums such as tablets, smartphones, and in virtual reality. Standalone or illusionary automated psychotherapy services can also be included in future research.

Further studies could identify, count, and scale behavioral cues. This requires manipulating non-verbal information when using technology. These cues, in general can aid the counseling pairs in psychotherapy [56], and from a clinical perspective, they can intensify the emotional reactions in psychotherapeutic communication [57, 58]. However, further research into the particular value of individual behavioral cues in CMP contexts is needed. This study did not isolate and compare these cues. Relatedly, in the present study, the role of emoticons and instant messaging features were excluded from the text-based mode of communication, but they may also be promising ways to improve written communication during CMP [59]. Including such features in future studies, either individually or combined, might provide further significant information on the rating of emotionality and the degree of emotional engagement that occurs via text.

It is also constructive for practitioners to be aware of the emotional recognition error rate. This error involves missed emotional cues and can be linked to bias in emotional perception [60]. Previous literature also suggests that emotional detection with high percentages of this error can result in lower rates of satisfactory care [60]. Thus, another requirement is to examine how satisfaction rates correlate with error rates in CMP communication.

The delivery of mental health care is an emotional endeavor. Therefore affective bias may have a major impact in CMP [61, 62]. A few studies have addressed how telecounseling affects the vicarious traumatization, or how much burnout in cybertherapists can be estimated [63, 64]. However, there has been no attempt to link indicators of the personal status of clinicians with remote psychotherapy. According to studies on emotional awareness, people who are aware of their emotional level identify more emotions in their partners. This factor is also relevant for psychotherapists who differentiate between emotional states of clients and assess the quality of the mental disorders daily [65, 66, and 67]. Based on these studies, there is a need to investigate such personal indicators in cybertherapists themselves because the therapeutic process can be affected.

Further research also needs to avoid cognitive biases. It has been well documented that clinicians may perform incorrect therapeutic processes due to dysfunctional thinking, problem solving, educational standpoint, clinical instruction, and training [25, 68]. Thus, the investigation of CMP should include the meta-cognition of cybertherapists such as their problem solving ability, memory, previous experiences, and their background knowledge. In the current study, heuristic decision-making and its link with training only was assessed. However, diagnostic process also has interconnection with awareness of clinicians whereas that ability was suggested a strategy to avoid cognitive biases in decisions [69, 70]. Therefore, more in-depth analysis of conscious, deliberate, slower decision making processes in CMP is needed.

Finally, the present study used self-report scales as a measure of the ability to perceive emotions in others. Future investigations should use more objective measures such as psychophysiological measurements, like EEG, fMRI, or GSR. These measures explore physical and cognitive activities of psychotherapists while assessing emotionality in different computer-mediated conditions [12,

71]. These objective data could then be compared with that collected from face-to-face psychotherapy sessions. Objective scales would provide information on cybertherapists' cognitions, which could be compared with self-report data.

5 CONCLUSIONS

The study describes the investigation of the emotional content of therapeutic communication via technology. In line with previous studies, it was thought that technological channels of communication can influence cybertherapists' ability to perceive emotions [60, 23, and 72]. Emotionality was introduced to examine the level of emotional perception via text, audio, and video modes of delivery. The method measured the perception of the intensity, valence, potency, and activation of the expressed emotions.

The perception of the valence and intensity of expressed emotions, as well as the potency and activation of emotional states, are sensitive to different computer-mediated delivery modes. That is, the perception of emotionality was lower in the text-based mode of therapeutic communication than in the audio-only and video-audio modes. It was also found that the robotic voice mode impeded participants' ability to identify expressed emotions. Therefore, cybertherapists need to consider the quality of the information they will receive via different modes of delivery (e.g., text messaging, telephone, videoconferencing). Analysis of the interview data showed that affective words, acoustic features, and bodily reactions could offer a range of important emotional cues, some of which may not be seen using computer-mediated delivery.

Furthermore, results showed a relationship between emotional perception and heuristic decision-making with computer-mediated communication. It was suggested therefore that cybertherapists can trust their emotional perception and subsequent diagnosis via technology; however, computer-mediated psychotherapy should still be provided with caution, particularly via emails or text messaging.

Finally, consistent with CMC literature, it has been shown that psychotherapeutic communication via technology can be influenced by missed nonverbal information, deviations in the interpretation of the content, variations in diagnostic process, and decreased levels in the reality of clients.

REFERENCES

- [1] Greenberg, L. S., & Paivio, S. C. (2003). *Working with emotions in psychotherapy* (Vol. 13). USA: Guilford Press.
- [2] Greenberg, L. (2004). Emotion-focused therapy. *Clinical Psychology & Psychotherapy*, 11(1), 3-16.
- [3] Greenberg, L. (2010). Emotion-focused therapy: A clinical synthesis. *FOCUS: The Journal of Lifelong Learning in Psychiatry*, 8(1), 32-42.
- [4] Ambady, N., & Rosenthal, R. (1998). Nonverbal Communication. In Friedman HS (Ed.), *Encyclopedia of Mental Health* (Vol. 2). New York: Academic Press.
- [5] Calvo, R. A., & D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *Affective Computing, IEEE Transactions on*, 1(1), 18-37. doi: 10.1109/T-AFFC.2010.1
- [6] Campbell, J., & Ehlert, U. (2012). Acute psychosocial stress: does the emotional stress response correspond with physiological responses? *Psychoneuroendocrinology*, 37(8), 1111-1134.
- [7] Edlund, B., & Sneed, N. (1989). Emotional responses to the diagnosis of cancer: age-related comparisons. *Oncology nursing forum*, 16(5), 691-697.
- [8] Spagnolli, A., & Bracken, C. C. (2011). Do You Feel As If You Are There? Measuring Presence in Cybertherapy. *In the proceedings of ISPR 2011 annual conference on the international society for presence research*, 2-9.
- [9] Walther, J., Van Der Heide, B., Ramirez, A. J., Burgoon, J. K., & Peña, J. (2015). Interpersonal and Hyperpersonal Dimensions of Computer-Mediated Communication. In S. S. Sundar (Ed.), *The Handbook of the Psychology of Communication Technology* (pp. 1-22). West Sussex, UK: John Wiley & Sons, Ltd.
- [10] Croes, E. A., Antheunis, M. L., Schouten, A. P., & Kraemer, E. J. (2016). Teasing apart the effect of visibility and physical co-presence to examine the effect of CMC on interpersonal attraction. *Computers in Human Behavior*, 55, 468-476.

- [11] Derks, D., Bos, A. E., & Grumbkow, J. V. (2008). Emoticons in computer-mediated communication: Social motives and social context. *CyberPsychology & Behavior*, 11(1), 99-101.
- [12] Frijda, N. H. (1986). *The emotions : Studies in emotion and social interaction*. Paris: Maison de Sciences de l'Homme.
- [13] Greenberg, L. (2008). Emotion and cognition in psychotherapy: The transforming power of affect. *Canadian Psychology/Psychologie canadienne*, 49(1), 49-59.
- [14] Stromer-Galley, J., Bryant, L., & Bimber, B. (2015). Context and Medium Matter: Expressing Disagreements Online and Face-to-Face in Political Deliberations. *Journal of Public Deliberation*, 11(1), article 1. Available at: <http://www.publicdeliberation.net/jpd/vol11/iss11/art11>.
- [15] Rochlen, A. B., Beretvas, S. N., & Zack, J. S. (2004). The Online and Face-to-Face Counseling Attitudes Scales: A Validation Study. *Measurement and Evaluation in Counseling and Development*, 37(2), 95.
- [16] Foley, G. N., & Gentile, J. P. (2010). Nonverbal communication in psychotherapy. *Psychiatry (Edmont)*, 7(6), 38-44.
- [17] Bauer, S., & Moessner, M. (2012). Technology-enhanced monitoring in psychotherapy and e-mental health. *Journal of Mental Health*, 21(4), 355-363.
- [18] Griffiths, K. M. (2013). A virtual mental health community—A future scenario. *Australian and New Zealand Journal of Psychiatry*, 47(2), 109-110.
- [19] Kang, S. (2007). Disembodiment in online social interaction: Impact of online chat on social support and psychosocial well-being. *CyberPsychology & Behavior*, 10(3), 475-477.
- [20] Walther, J. (1999). *Visual cues and computer-mediated communication: Don't look before you leap*. Paper presented at the Annual meeting of the International Communication Association, San Francisco, CA.
- [21] Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2010). *Nonverbal communication*. Boston, : MA: Allyn & Bacon.
- [22] Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social science information*, 44(4), 695-729.
- [23] Machado, P. P. P., Beutler, L. E., & Greenberg, L. (1999). Emotion recognition in psychotherapy: Impact of therapist level of experience and emotional awareness. *Journal of Clinical Psychology*, 55, 39-57. doi: 10.1002(SICI)1097-4679(199901)55:1<39::AID-JCLP4>3.0.CO2-V
- [24] Mishna, F., Levine, D., Bogo, M., & Van Wert, M. (2013). Cyber counselling: An innovative field education pilot project. *Social Work Education*, 32(4), 484-492.
- [25] Elstein, A. S., & Schwarz, A. (2002). Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. *British Medical Journal*, 324(7339), 729-732.
- [26] Klaczynski, P. A. (2001). Analytic and Heuristic Processing Influences on Adolescent Reasoning and Decision-Making. *Child development*, 72(3), 844-861.
- [27] Marewski, J. N., & Gigerenzer, G. (2012). Heuristic decision making in medicine. *Dialogues Clin Neurosci*, 14(1), 77-89.
- [28] Visschedijk, G. C., Lazonder, A. W., van der Hulst, A., Vink, N., & Leemkuil, H. (2013). Modelling human emotions for tactical decision-making games. *British Journal of Educational Technology*, 44(2), 197-207. doi: 10.1111/j.1467-8535.2012.01286.x
- [29] Bambling, M., King, R., Reid, W., & Wegner, K. (2008). Online counselling: The experience of counsellors providing synchronous single-session counselling to young people. *Counselling and Psychotherapy Research*, 8(2), 110-116.
- [30] Arbuthnott, K. D., Arbuthnott, D. W., & Thompson, V. A. (2006). *The mind in therapy: Cognitive science for practice*. Mahwah, NJ: Erlbaum.
- [31] Barrett, L. F. (2006). Valence is a basic building block of emotional life. *Journal of Research in Personality*, 40(1), 35-55.
- [32] Heise, D. R. (1989). Effects of emotion displays on social identification. *Social Psychology Quarterly*, 10-21.
- [33] Lively, K. J., & Heise, D. R. (2014). Emotions in affect control theory *Handbook of the Sociology of Emotions: Volume II* (pp. 51-75). Netherlands: Springer.
- [34] Niedenthal, P. M. (2008). Emotion concepts. *Handbook of emotions*, 587-600.
- [35] Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.
- [36] Hauke, G., & Dall'Occhio, M. (2013). Emotional Activation Therapy (EAT): Intense work with different emotions in a cognitive behavioral setting. *European Psychotherapy*, 1(11), 5-29.
- [37] Picard, R. W. (1997). *Affective computing* (Vol. 252). USA: MIT press Cambridge.
- [38] Walther, J., Loh, T., & Granka, L. (2005). Let me count the ways the interchange of verbal and nonverbal cues in computer-mediated and face-to-face affinity. *Journal of Language and Social Psychology*, 24(1), 36-65.
- [39] Chang Lee, K., Kwon, O., Kim, C.-R., & Kim, G. (2013). Factors affecting the intensity of emotional expressions in mobile communications. *Online Information Review*, 37(1), 114-131.
- [40] Bachorowski, J. A. (1999). Vocal expression and perception of emotion. *Current Directions in Psychological Science*, 8(2), 53-57.

- [41] Hancock, J. T., Landrigan, C., & Silver, C. (2007). Expressing emotion in text-based communication. *In proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 929-932. doi: 10.1145/1240624.1240764
- [42] Nass, C., Foehr, U., Brave, S., & Somoza, M. (2001). *The effects of emotion of voice in synthesized and recorded speech*. Paper presented at the proceedings of the AAAI symposium emotional and intelligent II: the tangled knot of social cognition, North Falmouth, MA.
- [43] Ortony, A., Clore, G. L., & Collins, A. (1990). *The cognitive structure of emotions*. USA: Cambridge university press.
- [44] Brockmeyer, T., Zimmermann, J., Kulesa, D., Hautzinger, M., Bents, H., Friederich, H.-C., . . . Backenstrass, M. (2015). Me, myself, and I: self-referent word use as an indicator of self-focused attention in relation to depression and anxiety. *Frontiers in psychology*, 6, Article 1564.
- [45] Zimmermann, J., Wolf, M., Bock, A., Peham, D., & Benecke, C. (2013). The way we refer to ourselves reflects how we relate to others: Associations between first-person pronoun use and interpersonal problems. *Journal of Research in Personality*, 47(3), 218-225.
- [46] El Ayadi, M., Kamel, M. S., & Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, 44(3), 572-587. doi: <http://dx.doi.org/10.1016/j.patcog.2010.09.020>
- [47] Vogt, T., André, E., & Wagner, J. (2008). Automatic recognition of emotions from speech: a review of the literature and recommendations for practical realisation. In C. Peter & R. Beale (Eds.), *Affect and Emotion in Human-Computer Interaction* (pp. 75-91). Berlin Heidelberg: Springer.
- [48] Siegman, A. W., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression: the effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology*, 102(3), 430-437.
- [49] Walther, J. (2011). Visual cues in computer-mediated communication: Sometimes less is more. In A.Kappas (Ed.), *Face-to-Face Communication Over the Internet: Issues, Research, Challenges*. (pp. 17-38). Cambridge , UK Cambridge University Press.
- [50] Nguyen, D. T., & Canny, J. (2009). More than face-to-face: empathy effects of video framing. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 423-432.
- [51] Dijksterhuis, A. (2004). Think different: the merits of unconscious thought in preference development and decision making. *Journal of Personality and Social Psychology*, 87(5), 586-598.
- [52] Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62, 451-482.
- [53] Martin-Khan, M., Flicker, L., Wootton, R., Loh, P.-K., Edwards, H., Varghese, P., . . . Gray, L. C. (2012). The diagnostic accuracy of telegeriatrics for the diagnosis of dementia via video conferencing. *Journal of the American Medical Directors Association*, 13(5), 487. e419-487. e424.
- [54] Sherko, E., Sotiri, E., & Lika, E. (2013). Therapeutic communication. *JAHN-European Journal of Bioethics*, 4(7), 457-466.
- [55] Boritz, T. Z., Bryntwick, E., Angus, L., Greenberg, L. S., & Constantino, M. J. (2014). Narrative and emotion process in psychotherapy: An empirical test of the Narrative-Emotion Process Coding System (NEPCS). *Psychotherapy Research*, 24(5), 594-607.
- [56] McKenna, K. Y., & Bargh, J. A. (2000). Plan 9 from cyberspace: The implications of the Internet for personality and social psychology. *Personality and social psychology review*, 4(1), 57-75.
- [57] Berenschot, F., van Aken, M. A., Hessels, C., de Castro, B. O., Pijl, Y., Montagne, B., & van Voorst, G. (2013). Facial emotion recognition in adolescents with personality pathology. *European child & adolescent psychiatry*, 1-8.
- [58] Minzenberg, M. J., Poole, J. H., & Vinogradov, S. (2006). Social-emotion recognition in borderline personality disorder. *Comprehensive Psychiatry*, 47(6), 468-474. doi: <http://dx.doi.org/10.1016/j.comppsy.2006.03.005>
- [59] Kotlyar, I., & Ariely, D. (2013). The effect of nonverbal cues on relationship formation. *Computers in Human Behavior*, 29(3), 544-551. doi: <http://dx.doi.org/10.1016/j.chb.2012.11.020>
- [60] Blanch-Hartigan, D. (2013). Patient satisfaction with physician errors in detecting and identifying patient emotion cues. *Patient Education and Counseling*, 93(1), 56-62.
- [61] Croskerry, P. (2005). Diagnostic Failure: A Cognitive and Affective Approach. In K. Henriksen, J. B. Battles, E. S. Marks, D. I. Lewin & P. Croskerry (Eds.), *Advances in Patient Safety: From Research to Implementation* (Vol. 2). Rockville (MD): Agency for Healthcare Research and Quality.
- [62] LeBlanc, V. R., McConnell, M. M., & Monteiro, S. D. (2015). Predictable chaos: a review of the effects of emotions on attention, memory and decision making. *Advances in Health Sciences Education*, 20(1), 265-282.
- [63] Furlonger, B., & Taylor, W. (2013). Supervision and the Management of Vicarious Traumatization Among Australian Telephone and Online Counsellors. *Australian Journal of Guidance and Counselling*, 1-13.
- [64] Taylor, W., & Furlonger, B. (2011). A Review of Vicarious Traumatization and Supervision Among Australian Telephone and Online Counsellors. *Australian Journal of Guidance and Counselling*, 21(02), 225-235.
- [65] Camodeca, M., & Rieffe, C. (2013). Validation of the Italian Emotion Awareness Questionnaire for children and adolescents. *European Journal of Developmental Psychology*, 10(3), 402-409.

- [66] Lahaye, M., Mikolajczak, M., Rieffe, C., Villanueva, L., Van Broeck, N., Bodart, E., & Luminet, O. (2011). Cross-validation of the Emotion Awareness Questionnaire for children in three populations. *Journal of psychoeducational assessment*, 29(5), 418-427.
- [67] Rieffe, C., Oosterveld, P., Miers, A. C., Meerum Terwogt, M., & Ly, V. (2008). Emotion awareness and internalising symptoms in children and adolescents: The Emotion Awareness Questionnaire revised. *Personality and Individual Differences*, 45(8), 756-761.
- [68] Norman, G. R., & Eva, K. W. (2010). Diagnostic error and clinical reasoning. *Medical education*, 44(1), 94-100.
- [69] Croskerry, P. (2003). The importance of cognitive errors in diagnosis and strategies to minimize them. *Academic Medicine*, 78(8), 775-780.
- [70] Hardy, M. A. (1993). *Regression with dummy variables (Quantitative applications in the social sciences)* (Vol. 93). London: AGE Publications India Pvt.Ltd
- [71] Sonnemans, J., & Frijda, N. H. (1994). The structure of subjective emotional intensity. *Cognition & Emotion*, 8(4), 329-350.
- [72] Mazzi, M. A., Bensing, J., Rimondini, M., Fletcher, I., van Vliet, L., Zimmermann, C., & Deveugele, M. (2013). How do lay people assess the quality of physicians' communicative responses to patients' emotional cues and concerns? An international multicentre study based on videotaped medical consultations. *Patient Education and Counseling*, 90(3), 347-353.
- [73] Reber, A., Reber, E., & Allen, R. (2009). *Dictionary of Psychology* (Vol. Fourth Edition). London, UK: Penguin Books.
- [74] Fiebler, R. (2002). How to do emotions with words: Emotionality in conversations. *The verbal communication of emotions*, 79-106.
- [75] Ray, W. J. (2012). *Evolutionary Psychology: Neuroscience Perspectives Concerning Human Behavior and Experience*. USA: SAGE Publications.
- [76] Pahl, K. (2009). Emotionality: A Brief Introduction. *MLN*, 124(3), 547-554.
- [77] Lang, P. J. (1995). The emotion probe: studies of motivation and attention. *American psychologist*, 50(5), 372.
- [78] Mather, M., & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, 6(2), 114-133.
- [79] Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual review of psychology*, 30(1), 527-554.
- [80] Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition and emotion*, 23(2), 209-237.
- [81] Cowie, R., & Cornelius, R. R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication*, 40(1), 5-32.
- [82] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. G. (2001). Emotion recognition in human-computer interaction. *Signal Processing Magazine, IEEE*, 18(1), 32-80.
- [83] Niedenthal, P. M. (2007). Embodying emotion. *science*, 316(5827), 1002-1005.

© 2017 by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

APPENDIX next page

APPENDIX

There is a problem with how people describe emotionality in technology-based communication. Emotionality is a measure of a person's emotional reactivity or response to a stimulus, and is relevant in the transmission and comprehension of communication [43, 73]. Based on existing theory and research [74, 75], it could be said that emotionality is defined as labeling emotions, and articulating different emotional experiences. In psychotherapy, however, describing emotionality requires perceiving and explaining the quality of the dominant emotional feeling, rather than exploring isolated emotions [76]. In other words, emotionality measures the qualitative, dynamic, interactive, and emergent indexes of a client's emotional expressions. As yet, no conclusive study linking emotionality to CMP has been carried out.

Therefore, the Emotionality in Cyberspace (EiC) instrument was developed by the author to evaluate a person's ability to perceive the emotional content of CMP sessions. The instrument was based on other studies that were designed to assess ability to detect emotions, or emotional expressions in face-to-face interventions [13,1]. It should be noted that different methods are used to measure ability to detect emotional states in others. One major method is that favored by neuropsychologists [77, 43, and 37]. This method uses a two-dimensional space. The valence of emotion is the first dimension and represents the negative and positive experiences of emotion in people². The second dimension is the intensity of emotions, or the arousal of a person when encountering an event, object, or action [12, 77, and 37]³.

Another method for studying ability to perceive emotions, which is of particular interest for the current project, suggests that it should be possible to infer a person's emotional state from their behavior (e.g., vocal, facial, and body characteristics) [79, 12]. This method appears to support the two-dimensional space method for the assessment of behaviors [72]. Arbuthnott, Arbuthnott, and Thompson [30] suggests, for example, that higher levels of pitch and amplitude in conversation are associated with higher levels of emotional arousal and valence of speakers [80].

However, the measure developed for the current study suggests a four-dimensional space for emotional perception. In addition to 'valence' and 'intensity', which are the main dimensions within psycho-communicative contexts, 'potency' and 'activation' of emotions have been added as other dimensions of emotional communication. The dimensions, referred to hereafter as emotional factors of communication, are introduced below.

Emotional valence: Valence relates to the quality of emotional arousal, attraction or aversion to people, things or events during an interaction between speakers [31]. Therefore, it aids in the process of evaluating whether another person's feelings are helpful or risky. This dimension of emotional communication generally refers to a capacity to discern between the positive or negative tone of affective expressions [30].

Emotional Intensity: Intensity refers to how individuals express their emotions. This attribute, which can result in a range of behavioral outbursts (e.g., emphasized affective words, crying, yelling, or physical expressions), indicates the perceived magnitude of the emotions expressed [30].

Emotional Potency: Affect Control Theory describes potency as a person's level of "powerlessness" to "powerfulness" in certain situations [32, 33]. In other words, potency is the capacity to produce strong emotional reactions in perceivers. For instance, a nonverbal cue

² This dimension is linked to heart rates and both right and left frontal lobes [30].

³The intensity level of the emotional state is represented by cardiovascular and respiratory activities including heart rate and the electrical conductance of the skin due to eccrine (sweat) gland activity [80].

demonstrating a potent emotion would be rotating one's body away from the screen; others might be consistent gaze behaviors (e.g., staring), facial expressions (e.g., lowered brows), and vocal changes (e.g., amplitude). Affective words also convey emotional potency [34]. Based on Osgood's semantic analyses [35], high potency is characterized by controlled and deliberate articulation of affective words, while low potency or powerlessness is articulated by uncontrolled, unstable, and loose characteristics. The characterizations of the controlled side of potency include highlighting direct or indirect affective words, or affixing a label to the words such as positive adjectives (big, powerful, deep, strong, high, long, full, and many) or negative adjectives (little, powerless, shallow, weak, low, short, empty, and few).

Emotional Activation: Emotional activation describes a direct emotional reaction in a client's behavior and indicates how the client is, consciously and cognitively, connected to his or her own feelings [36]. Emotional activation, which is a significant component of Emotional Activation Therapy (EAT), shows the strength of the person's disposition to take some action in response to one or more emotions [81, 82]. This factor, like potency, is represented via direct emotional words, low or high rate of facial expressions, bodily movement, and speech [83, 34, and 35].

To assess how emotional states are perceived in communication, participants in the study rated the above factors between '0' (not at all) to '100' (a great deal).

The EiC instrument also included one closed and two open-ended questions. To test participants' decision-making ability, the closed question asked whether or not they believed that the person in each scenario had a mental health issue. Possible responses were 'no problem', 'personal issues but not significant', and 'significant issue(s) that need professional help'. The first open-ended question asked participants to describe or name the emotions they observed in each condition of computer-mediated communication (CMC). The second open-ended question asked participants to explain the types of informational markers (e.g., affective words, vocal or facial cues) they observed in the scenarios.

Measuring Emotionality in Therapeutic Communication

Counseling pairs in psychotherapy sessions share different personal, social and mental experiences, and as communication runs, the content of these experiences can change or the urgency of them decreases. Therefore, psychotherapists do not count on one feeling or detached emotion (e.g., sadness) when they deal with mental disorders such as depression. Rather, psychotherapists analyze the emotional content across time in a continuum of therapeutic communication [74, 76].

So far, there has been no attempt to calculate emotionality when it occurs in communication across time. In this study, a comprehensive formulation of emotional perception is introduced. By using the following formula, the quantity of a clinician's perception of emotional communication is estimated. To calculate emotionality, the valence (v), intensity (i), potency (p) and activation (a) factors are integrated in a linear model⁴. Thus, emotionality is calculated by:

$$E_c = \sum R(v, i, p, a)$$

where E is the number of perceived emotions and c represents the conditions through which E was rated.

Further, based on a study by Machado et al. (1999), differences in the clinicians' judgments of the emotional content were defined as an 'emotional deviation score' [23]. This definition provides

⁴It is noted that the current measure of emotionality disregards the type of emotions (e.g., happiness or sadness) that might be identified in the stream of therapeutic communication.

us, first, with a way to calculate how much emotional content is observed, and second, differences between experts' and non-experts' observations can be used to establish if there are any differences in the perceived level of emotionality transmitted via different technologies. The following formula for the emotional deviation score is represented by:

$$\Delta E_c = \sum (nonexperts(E_c) - experts(E_c))$$

where ΔE is the difference between the emotionality observed by non-experts and experts. However, unlike the definition of Machado, ΔE in the current definition is calculated without squaring the differences of participants' ratings. This was done so that the polarity of the ratings, so to speak, a negative outcome in the emotional deviation score, ΔE , will indicate an inability to perceive the emotions expressed by the speakers and thus, in turn, negativity in the emotional content of communication will be noticed.

To test if the model was dependable, participants rated the variables (e.g., intensity_ separately for each mode of delivery. Emotionality, E_c , and emotional deviation scores, ΔE_c , were calculated afterward. Then, a test of the internal consistency was conducted on the data. Cronbach's coefficient for emotionality scores was 0.75 in Scenario A (representing a psychological problem) and 0.60 in Scenario B (a discussion about a mental health problem). Based on Cronbach's values, the internal consistency of emotional deviation scores was also checked to examine the goodness of fit, which was remarkably high for Scenario A (i.e., 0.81), while it was respectable for Scenario B (i.e., 0.67).

Emotionality in Cyberspace (EiC) Questionnaire

You will experience a version or 2 of the same event/situation: text-alone, audio-alone, vision-alone, audio+vision. I am interested in which aspects, if any, you think have emotion content or convey emotion, and how strong that emotion is. (next page)

Participant: _____

Age:

Gender: Male Female Other

How would you classify yourself?
 Native English Speaker
 Non-Native English Speaker

Education level:
 Secondary school
 Diploma
 Under graduate
 Postgraduate
 Other

Expertise:
 Psychology/counselling
 Medicine/Nursing
 Social/Human Sciences
 Engineering/Natural Sciences
 Other

Stimulus: _____ Place in order: _____

1. On the scale below please indicate how much emotional valence (positive or negative) you think the vignette conveys

0|-----|-----|-----|-----|-----|-----|-----|-----|-----|100
None Some Moderate Quite a bit A Lot

2. On the scale below please indicate how intense the emotion in the vignette is

0|-----|-----|-----|-----|-----|-----|-----|-----|-----|100
None Mild Moderate Strong Severe

3. On the scale below please indicate how potent you think the emotion in the vignette is (i.e., do you feel effected by it?)

0|-----|-----|-----|-----|-----|-----|-----|-----|-----|100
None Mild Moderate Strong Severe

4. On the scale below please indicate how active you think the emotion in the vignette is

0|-----|-----|-----|-----|-----|-----|-----|-----|-----|100
None A little Somewhat Quite a bit Severe

5. Do you think the emotion conveyed is real or acted?

Real
 Acted

6. Do you think the vignette conveys a psychological problem?

No problem
 Personal issues but not significant
 Significant issue(s) that need professional help

Interview Questions:

7. Please name/describe the emotion(s) you perceive in the vignette.