Automatic Mimicry Detection in Medical Consultations

Kaihang Wu, Chunfeng Liu, Silas Taylor, Paul W.B. Atkins, Rafael A. Calvo *Snr Member, IEEE*

***Abstract*— Good communication and rapport between patients and doctors is important to achieve positive health outcomes. Mimicry, when a person copies the behavior of someone else, has been related to good communication and building rapport. We develop a method to automatically detect non-verbal mimicry and study the relationship between mimicry and the perceived quality of the interaction. Automating the detection of mimicry is important since it may be a proxy for empathy and perspective taking, both features of high quality interactions for caring professions. This pilot study involved 91 sessions between medical students and volunteers acting as patients. The results confirm a correlation between our measure of mimicry and the assessed quality of the interaction. Therefore, medical students may improve their communication skills by practicing behavioural mimicry.**

I. INTRODUCTION

Medical consultations play an important role in clinical care. During these consultations clinicians learn the information necessary to make a diagnosis, patients learn about their illness, and when necessary, clinicians recommend a therapy [1]. Medical communication skills are acknowledged as one of the most important factors to achieving healthier outcomes [2].

Because doctor-patient communication is so important, a body of research in health sciences and medicine has focused on identifying the communication styles that are most successful. Both verbal and non-verbal features can be studied, but most research has focused on the verbal [3].

Here we focus on non-verbal features broadly classified into four groups: (1) head and body movements, including facial expressions and gestures; (2) vocalics, such as speech rate and pitch; (3) body contact, such as a handshake, (4) proxemics, for instance, how far a doctor is away from the patient or which direction is the doctor's body toward [3].

The study of such features is important as shown by reports that smiles, frowning, head nodding and shaking are correlated to certain medical outcomes [4]. [5] showed evidence that more smiles and nodding may be perceived by patients as more empathy and rapport.

Rapport and sense of empathy are considered important in effective therapeutic communication [5]. Furthermore, there is evidence of a positive connection between mimicry and rapport [6–8]. Although rapport is a complex social behaviour that can be interpreted through various social cues, nonverbal mimicry is a reliable indicator because it can cultivate similarity and express understanding between the parties [8]. A few studies on doctor-patient communications highlight the importance of rapport in medical consultations. For example, [9] found that clinicians' automated mimicking builds mutual affect during the interaction, and this links to one criteria of rapport i.e. concordance. From this body of evidence, we hypothesize that rapport is correlated with mimicry.

Previous studies have manually annotated nonverbal mimicry. For example, a study adopted 4 raters to code behavioural mimicry in videotaped conversations [10]. An instance of behavioural mimicry happens when the behaviour of the mimic matched the action of the other within a 10-second window. In another study, mother-infant interactions were recorded by two individual cameras and coded by 16 untrained raters to judge the behaviours of mother and baby following a rating form in the re-created dyadic interaction (combine two videos into one with split screens) [11].

There is a dearth of research on how to automatically annotate instances of mimicry. In [12], this was done by calculating cross-correlations of pixel intensity in accumulated motion images of two persons (within 20 second windows). A systematic review of this literature has not been done, but to our knowledge none have used mimicry evaluations in the context of medical communication skills training.

This paper contributes a technique for automatically identifying mimicry in dyadic interactions. In addition, it explores the relationship between students' nonverbal mimicry and their overall communication skill during tele-consultations.

II. METHOD

We conducted a 4-month user study between undergraduate students and actor Simulated Patient (SP). Students were required to complete one tele-consultation with a SP through the EQClinic platform [13]. In addition,

*RAC is supported by an Australian Research Council Future Fellowship FT140100824

Kaihang Wu, is with the Electrical and Information Engineering, University of Sydney, Camperdown, NSW 2006 Australia (phone: 61401595094 e-mail: kaihang.wu@sydney.edu.au)

Chunfen Liu is with the Electrical and Information Engineering, University of Sydney, Camperdown, NSW 2006 Australia (chunfeng.liu@sydney.edu.au)

Silas Taylor is with UNSW Medicine, University of New South Wales, Ground Floor, Wallace Wurth Building (C27), Botany St, Randwick

Paul WB Atkins is with the Institute for Positive Psychology and Education at the Australian Catholic University. 25A Barker Road, Strathfield NSW 2135 Australia

Rafael A Calvo is with the Electrical and Information Engineering, University of Sydney, Camperdown, NSW 2006 Australia (Rafael.calvo@sydney.edu.au)

EQClinic allows the SP to assess the performance of students and provide different forms of feedback.

A. Participants

Participants were 91 Year-2 medical students from an Australian university and 2 SP (SP1: female, SP2: male). All students were enrolled in a communication skills training course provided by the medical school. The two SPs were experienced actors and were recruited by a job advertisement website. Neither students or SPs had experience in tele-consultations using EQClinic. This study was approved by the UNSW Research Ethics Committee (HC Reference Number: HC16048).

B. Instruments

The Student-Patient Observed Communication Assessment (SOCA), an adapted version of the Calgary Cambridge Guides [14], was used in the study to measure students' communication skills. SOCA covered four aspects: providing structure, gathering information, building rapport, and understanding the patient's needs. Each item was rated on a four-point scale.

C. Procedures

Tele-consultation procedure

Students interviewed SPs through the tele-conference component of EQClinic, and all interviews were recorded. After the interview, the SP assessed the performance of the student using SOCA and students could review their SOCA results.

Data collection

55 students completed a tele-consultation with SP1 and 36 students completed a tele-consultation with SP2. Two separate video recordings were generated for each consultation by the platform: one for the SP and one for the student. The video recordings had 25fps frame rate and resolutions of 640*480 pixels with the format of WebM.

Figure 1: System Architecture

Fig. 1 illustrates the architecture of the whole system for this study. EQClinc platform as the first part of the system was to collect raw data in terms of nonverbal behaviours of medical students and SPs. The second step was to get the time periods of four nonverbal behaviours respectively. For each interview, the periods of nonverbal behaviours were collected from both student's and SP's video recordings. Those periods were for the next part of the system, mimicry detection. According to the definition of mimicry, the nonverbal mimicry got detected by computing and analysing the time difference between the time slots of nonverbal behaviour belonging to SP and student. At the end, we obtained frequency of occurrences of nonverbal mimicry during an interview.

To detect mimicry within consultations, four kinds of nonverbal behaviour were automatically identified: smile, frown, nodding and head shaking, for both students and SPs.

Nonverbal behaviour detection

Recordings of each session were pre-processed to identify nonverbal behaviour features. For smile and frown detection, we used a facial expression detection library provided by Affectiva. This library detects the smile or frown intensities (values between 0-100) of each frame of the video recordings. The smile and frown intensity values along with the frame numbers were stored in two separate text files.

The algorithm proposed by [15] was adapted to detect nodding and shaking. Instead of observing the movement of the "mid-eye" point, the movement of the nose point was used.

Periods of behaviours

Smile and frown intensities were measured for each frame, but what we actually need was the smiling and frowning periods. In order to eliminate spurious signals, the behaviours were considered active only in periods when the intensity value was greater than 50. Otherwise they were considered inactive. If a behaviour in a frame was active and continued to be active in the next, then the algorithm assumed that the current frame was the starting point and the last positive frame of the sequential frames was the endpoint of the period. Since video was recorded at a fixed rate of 25 fps, the number of frames can be converted to time. In the next step, we merged the time periods. If the interval between two periods was shorter than 1 second, these two periods were considered as one longer period for this nonverbal behaviour.

Because the algorithm for head nodding and shaking detection yield the starting time and ending time of nodding and shaking as the result, there was no further processing on these two behaviours.

Data Analysis Procedure

Mimicry definition

We consider that the student's nonverbal behaviour mimicked the SP's behaviour if one of four following conditions was presented：1) the starting point of medical student's behaviour happened after the SP's and the ending point of student's behaviour was earlier than the ending point of the SP's; 2) the starting and ending points of the student's behaviour were both later than the starting and ending points of the SP's; 3) the starting point of the students' behaviour was at the ending point of the SP's; 4) the starting point of the student's behaviour happened after the SP's within a 3-second window.

Mimicry detection

Based on the above conditions, a method was developed to automatically detect mimicry. The main idea of the method was that, if the current nonverbal slot of SP and a late period from medical student was in the one of four conditions, for instance, if a student smiled after the SP's smiling 1 second later, this was considered an instance of mimicry. Otherwise, the algorithm inferred that no mimicry had occurred.

Correlation calculation

To investigate the effects of nonverbal mimicry on assessed communications skills we selected the 10 students with highest and the 10 with the lowest SOCA scores, in the consultations with both SP1 and SP2. The video recordings of these 40 students were analysed to explore the correlations between assessed communication skill and nonverbal mimicry.

In this study, the main measurement of student's performance was SOCA that included an assessment of rapport. The maximum scores for SOCA and rapport were 20 and 4 respectively.

III. Results

Table 1 reports average times of four nonverbal mimicry behaviours, the mean SOCA scores and mean rapport scores of different student groups. For example, in SP1's sessions, the average times of detected smiling mimicry was 13.10 for low-score students and 16.20 for high-score students. The mean frequency of occurrences of smiling mimicry for SP2's sessions was 3.56 for low-score students, and 8.80 for high-score students. Similarly, we also computed average frequency for mimicry of frowns, head nodding and head shaking of all the groups. Rated by SP1, low-score group got the average overall SOCA 6.70, however, the mean score of high-mark group was 13.90. And in SP2's assessments, the mean score of low-mark students was 7.78 but that of high-score students was 13.50. Scores of building rapport were calculated the averages as well for the further analysis.

In Table 2, a Pearson product-moment correlation coefficient was computed to assess the relationship between the frequency of mimicry and overall rating of SOCA and rapport scores respectively. The frequency of head shaking mimicry with SOCA ($r = 0.46$, $p < 0.05$) and rapport scores given by SP2 ($r = 0.54$, $p = 0.01$) were both significantly correlated. Rapport scores were also positive correlated with nodding with high significant level ($p < 0.05$) in SP2's session. But, there was a nonsignificant correlation between the nodding mimicry and SOCA scores.

IV. Discussion

We examined the relations between medical students' nonverbal mimicry and simulated patient assessments of their communication skill. The first result suggests that high-score students' mimicry in smiles, frowns, head nodding and head shaking were more than those of lower-score students for all nonverbal behaviours with both SPs (except frown with SP1). The differences of frequency of appears between high-score students group and low-score students group support our hypothesis that communication skills might be associated with nonverbal mimicry.

Table 2. Correlation between times of mimicry of nonverbal behaviours and SOCA, rapport scores

		Correlations (with significant levels p)			
		smile	**frown**	**nodding**	**shaking**
SP1(SOCA)	r	0.26	-0.20	0.37	-0.18
	p	0.28	0.40	0.11	0.45
SP1(rapport)	r	0.38	-0.21	0.33	-0.13
	p	0.09	0.39	0.16	0.57
SP2(SOCA)	r	0.31	-0.25	0.40	0.46*
	p	0.19	0.29	0.08	0.04
SP2(rapport)	r	0.43	-0.22	0.47*	0.54**
	p	0.06	0.34	0.04	0.01

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed)

Results of correlation provide interesting results on how the amount, and type, of nonverbal mimicry relates to SOCA and rapport scores. For example, for SP1 (female), the mimicry of smile and nodding from participant students might be evidence of their communication skills. According to the results, smiling mimicry had positively low-significant correlation with student's rapport ($p = 0.09$ n.s), and a nonsignificant correlation (0.11) existed between nodding mimicry and overall SOCA scores. But, there was no correlation between neither the mimicry of frown nor shaking and students' interaction skill in SP1's perception. It could be that nonverbal mimicry of negative behaviours can impulse counterpart's negative emotions which is not expected.

For SP2 (male), there seems to be a strong relationship for shaking. It was clearly positively associated with both overall communication assessment and rapport. However, that was completely different with the result of SP1. Gender differences may be a reason as suggested in the literature [16]. Students' mimicry of smiling might positively influence their interacting skills because the correlation value was 0.43 and p value is 0.06. Also, mimicry of nodding can be another

Table 1. Average times of mimicry of nonverbal behaviours with average scores of SOCA and rapport of students for different SPs (SD: standard deviation)

	smile		**frown**		**nodding**		**shaking**		**Overall SOCA**		**rapport**	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
SP1(low)	13.10	5.38	0.90	1.31	13.10	5.82	3.80	1.75	6.70	1.54	1.70	0.44
SP1(high)	16.20	7.41	0.20	0.57	16.20	8.11	4.20	2.25	13.90	1.31	3.60	0.47
SP2(low)	3.56	1.81	3.22	3.95	3.56	1.13	3.78	2.01	7.78	1.81	2.22	0.61
SP2(high)	8.80	6.95	3.90	4.70	8.80	2.60	10.40	5.69	13.50	1.37	3.50	0.48

indicator of good communication skills because of the nonsignificant levels (p_{SOCA} = 0.08, $p_{rapport}$ = 0.04).

V. Conclusion and Future Work

This study described a method for automatically annotating nonverbal behavioural mimicry in video interactions. Computer vision techniques that produce such automated annotations can be used to improve human-human communications. This is important in many domains where empathy and rapport can increase the likelihood of behaviour change and wellbeing, such as in the case of doctor-patient interactions.

This pilot study provides early evidence that medical students' nonverbal mimicry is correlated with their communication skills. Future studies will explore this relationship in a larger group of students with varying skills. This mimicry detection technique will also be used to study how different physician nonverbal behaviours are related to patient satisfaction in different gender and cultural background combinations.

References

[1] R. Iedema and M. Manidis, "Patient-clinician communication: An overview of relevant research and policy literatures," 2013.

[2] R. Street, G. Makoul, N. Arora, and R. Epstein, "How does communication heal? Pathways linking clinician–patient communication to health outcomes," *Patient Educ.*, 2009.

[3] C. Liu, K. Scott, and R. Calvo, "Towards a reflective online clinic (ROC): a tool to support interviewing skills," *CHI'15 Work. Dev. Ski. Wellbeing*, 2015.

[4] J. Crane and F. G. Crane, "Optimal nonverbal communications strategies physicians should engage in to promote positive clinical outcomes," *Health Mark. Q.*, vol. 27, no. 3, pp. 262–274, 2010.

[5] D. Buyck and F. Lang, "Teaching medical communication skills: a call for greater uniformity," *Fam. Med. CITY-*, vol. 34, no. 5, pp. 337–343, 2002.

[6] J. L. Hagad, R. Legaspi, M. Numao, and M. Suarez, "Predicting levels of rapport in dyadic interactions through automatic detection of posture and posture congruence," in *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, 2011, pp. 613–616.

[7] J. N. Bailenson, N. Yee, K. Patel, and A. C. Beall, "Detecting digital chameleons," *Comput. Human Behav.*, vol. 24, no. 1, pp. 66–87, 2008.

[8] S. Bilakhia, S. Petridis, and M. Pantic, "Audiovisual detection of behavioural mimicry," in *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, 2013, pp. 123–128.

[9] A. Finset and T. A. Mjaaland, "The medical consultation viewed as a value chain: a neurobehavioral approach to emotion regulation in doctor–patient interaction," *Patient Educ. Couns.*, vol. 74, no. 3, pp. 323–330, 2009.

[10] M. Stel and R. Vonk, "Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction," *Br. J. Psychol.*, vol. 101, no. 2, pp. 311–323, 2010.

[11] F. J. Bernieri, J. S. Reznick, and R. Rosenthal, "Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions.," *J. Pers. Soc. Psychol.*, vol. 54, no. 2, p. 243, 1988.

[12] X. Sun, K. P. Truong, M. Pantic, and A. Nijholt, "Towards visual and vocal mimicry recognition in human-human interactions," in *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*, 2011, pp. 367–373.

[13] C. Liu, K. M. Scott, R. L. Lim, S. Taylor, and R. A. Calvo, "EQClinic: a platform for learning communication skills in clinical consultations," *Med. Educ. Online*, vol. 21, no. 1, p. 31801, 2016.

[14] A. Simmenroth-Nayda, S. Heinemann, C. Nolte, T. Fischer, and W. Himmel, "Psychometric properties of the Calgary Cambridge guides to assess communication skills of undergraduate medical students," *Int. J. Med. Educ.*, vol. 5, p. 212, 2014.

[15] S. Kawato and J. Ohya, "Real-time detection of nodding and head-shaking by directly detecting and tracking the' between-eyes,'" in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, 2000, pp. 40–45.

[16] M. S. Mast, "On the importance of nonverbal communication in the physician–patient interaction," *Patient Educ. Couns.*, vol. 67, no. 3, pp. 315–318, 2007.