The What, When, and Why of Facial Expressions: An Objective Analysis of Conversational Skills in Speed-Dating Videos

Mohammad Rafayet Ali†, Taylan Sen†, Dev Crasta*
†Department of Computer Science, *Clinical and Social Science in Psychology
University of Rochester, United States

Viet-Duy Nguyen†, Ronald Rogge*, Mohammed (Ehsan) Hoque†
†Department of Computer Science, *Clinical and Social Science in Psychology
University of Rochester, United States

Abstract—**In this paper, we demonstrate the importance of combinations of facial expressions and their timing, in explaining a person's conversational skills in a series of brief non-romantic conversations. Video recordings of 365 four-minute conversations before and after a randomized intervention were analyzed in which facial action units (AUs) were examined over different time segments. Male subjects (N=47) were evaluated in their conversation skills using the Conversational Skills Rating Scale (CSRS). A linear regression model was used to compare the importance of AU features from different time segments in predicting CSRS ratings. In the first minute of conversation, CSRS ratings were best predicted by activity levels in action units associated with speaking (Lips part, AU25). In the last minute of conversation, affective indicators associated with expressions of laughter (Jaw Drop, AU26) and warmth (Happy faces) emerged as the most important. These findings suggest that feedback on nonverbal skills must dynamically account for shifting goals of conversation.**

***Keywords-** speed-date; conversational skills; facial expressions;*

I. Introduction

For the typical individual, navigating our social society requires several social interactions per day, from brief mundane business transactions to longer intimate conversations in our personal relationships. While these interactions have varying functions, throughout these unstructured contexts, people wish to be seen as socially effective. This need becomes especially important when meeting new people and forming new social relationships.

Common wisdom suggests a number of "simple" pieces of advice to keep in mind in such meetings such as to smile, speak clearly, and make eye contact. However, rigid adherence to these rules can be ineffective: non-stop smiling is likely to give the impression of being insincere while unbroken eye-contact can feel disconcerting. Prior work has argued for incorporating the timing of facial expressions in understanding their significance [1]. Popular advice suggests that absolute first impression (i.e., the first few seconds to one minute of a conversation) may be more important than the rest of an interaction [2]. In contrast, there are many popular conceptions that emphasize final impressions (e.g., "ending on a good note;" "leaving on good terms"). Is it possible to determine which of these notions hold true in an objective way? Is there any valid general advice we can give in order to improve global conversational skills?

These questions are especially important for individuals with social anxiety challenges or autism spectrum diagnoses. Such individuals might feel overwhelmed by the high volume of often conflicting advice, and could benefit from objective direction and feedback about their social skills. While individuals may be able to gain this direction from a human coach, such coaching would be very costly. Additionally, the very stigma around social skills may make one reluctant to seek out such help even when available. An automated conversation skills evaluation tool could help users improve their conversation skills by identifying specific nonverbal behaviors to practice at their own convenience in the privacy of their own home.

In this study, we try to address the above questions to develop a generalizable understanding of the characteristics of "effective communication." More specifically, we present a novel automated methodology for evaluating conversational skills and identifying several facial action unit features associated with good conversational skills use in a friendly encounter with a stranger. The analyzed data includes video recordings of 365 four-minute interactions between 47 male individuals rotating between 8 trained female research assistants in a series of brief non-romantic conversations using a speed-dating structure [3]. The conversational skills of each of the male participants were evaluated using several metrics from the Conversational Skills Rating Scale (CSRS) [4]. Linear regression and support vector machine models were used for automatically predicting communication skill ratings from features extracted from the videos.

In this paper we attempt to answer the following research questions:

- Is it possible to predict conversational skills ratings through use of automated facial expression analysis tools?
- Can we observe which facial expression combinations are effective, and when they might be the most beneficial over the course of a conversation with a new acquaintance?

978-1-5386-2335-0/18/$31.00 ©2018 IEEE

In summary, we find that understanding human coders' ratings of conversation skills across multiple speed-dating interactions is not an intractable problem. Instead, average conversational skill ratings can be reasonably well explained (r = 0.68) using affective visual features with a linear regression model. We identified that the features associated with happiness (AU6 - cheek raiser, AU12 - lip corner puller, AU14 – dimpler) and speaking (AU25 - lips part, AU26 - jaw drop) obtained during the first and last minute of the conversation were most important in predicting conversational skill ratings.

II. Background

While there have been numerous studies examining automated analysis of face-to-face interactions [5], only a handful of studies provide tools to help speakers in evaluative contexts. Even among these studies, the prediction is often social situations with very specific goals such as public speaking [6] and job interviews [7] [8]. However, it is unclear whether these findings can be extended to more common social settings where individuals do not simply aim to impress an evaluator but be liked by the evaluator.

One effective tool for studying interpersonal attraction has been the speed-dating methodology, which allows researchers to examine social behaviors demonstrated across several interactions [9] [10]. Veenstra and Hung [11] used automated analysis of the position, proximity, and motion from overhead videos to predict a speed-daters' interest in their partner. Similarly, Ranganath et al. [12], [13] used automated nonverbal speech analysis to detect attempts to flirt. In both cases, the nonverbal analysis effectively predicted a given speaker's feelings towards their partner. However, effective prediction of the partner's interest in the speaker has only been achieved through manual coding of facial expressions and posture [14]. However, it is this latter prediction that is most critical, as users do not need feedback on their own feelings (which are known to them) but aim to know how others will respond to their behavior. The present study attempts to address this gap through an automated analysis of affective visual features expressed during a speed-dating interaction.

Reviewing the literature on prediction of others' evaluations suggests multiple possible approaches to utilizing a large amount of data that can be extracted from an extended interaction to predict a single evaluation given at the end of the interaction. The most common approach is to calculate a summary statistic of a given indicator across the whole sample (e.g., using a total number of pauses across an interaction [6], [8], [14]). However, the social psychological literature has demonstrated that relatively small samples of a behavior contain enough information for human observers to predict others' impressions [15], [16]. This "thin-slicing" approach can be further complicated by the choice of which moment in time is analyzed. When Carney et al. [1] investigated the accuracy of judges' first impressions in predicting several personality traits, they found that the specific location of the sample – at the beginning, middle, or end of an interaction – impacts its predictive utility. As the ability to draw conclusions from small samples of an interaction would have notable implications for how rapidly an automated coach can offer users feedback, this study compared the predictive utility of separate points of an interaction to the interaction as a whole and to one another.

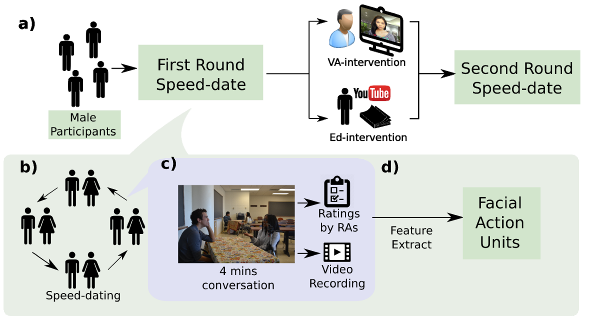


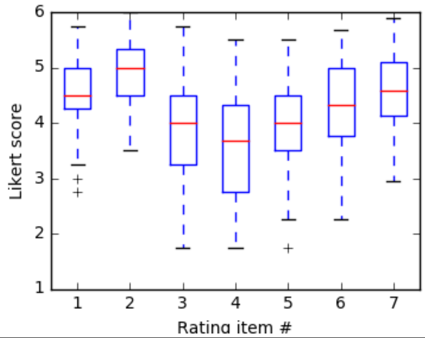
Figure. 1. Study design a) two rounds and two randomized intervention arms, b) male participants were rotated among female RAs within each round, c) each speed-date is a 4 minute conversation which results in a video recording and a rating evaluation, d) facial action units were extracted from the videos for analysis

III. Study Design

The raw data used in this paper was gathered from a previously published speed-dating study [3] which is shown in Figure 1. In summary, 365 four-minute speed-dates were recorded between 47 male participants and 8 female research assistants ("RAs"), after obtaining the participants' permission. All participants were native English-speaking undergraduate students aged between 18 and 23 years old. All of the RAs were psychology students and were extensively trained to be consistently friendly and encouraging across interactions. Each participant had the first round of speed-dates with 3 to 5 different RAs, followed by random assignment to one of two interventions. The intervention received by the male participant was one either: 1) an interaction with a virtual agent coaching system ("VA-intervention", N=23), or 2) an educational intervention consisting of an instructional video [17] and an informational pamphlet [18] on improving conversation skills ("Ed-intervention", N=24). Both the virtual agent and the materials in the educational intervention specifically targeted four nonverbal features: smiling, volume, body movement, and eye contact. After the intervention, each participant then completed a second round of speed-dates with the same RAs he had in the first round. The RAs were blind to which intervention the male participants received.

In the original study, the resulting data allowed us to combine the strength of blinded pre-post ratings from treatment evaluations with the ability to isolate speaker effects across repeated conversations of speed-dating designs. In the present study, the blind ratings mean that both the first and second speed-dating rounds provide two separate opportunities to assess the relationship between specific affective features of interaction videos and overall ratings.

TABLE I. LIST OF RATING ITEMS AND DISTRIBUTION

Name	Item Description	mean	std	Distribution box plot
Overall Effectiveness	Composite: Skilled/Competent/Effective Communicator; Good Conversationalist;	4.64	0.68	
Overall Appropriate	CSRS Molar: Appropriate Communicator	4.98	0.54	
Smiling	CSRS: Smiling and/or laughing	3.97	0.99	
Humor	CSRS: Use of humor and/or stories	3.70	0.97	
Vocal	CSRS: Vocal Confidence	4.02	0.76	
EyeCon	CSRS Use of Eye Contact	4.29	0.74	
Likeability	Average of 7 interpersonal attractiveness items	4.63	0.67	

IV. DATA COLLECTED

A. *Ratings of Participant Conversational Skills*

Table I list the individual and combined ratings used for our analysis. Each of the RAs was trained to rate participants on selected items from the Conversational Skills Rating Scale (CSRS) prior to the study. The CSRS is a well-validated tool to help raters provide reliable, relatively objective assessments of an individual's conversation skills [4]. To provide the most accurate assessment, RAs immediately rated the participants after each interaction of a speed-dating round. First, RAs provided quick global assessments of using the CSRS' "molar" items, which ask raters to evaluate the participants' overall performance. Due to their high intercorrelations, four of the CSRS molar items measures were averaged into the Overall Effectiveness measure (Cronbach's α=0.929). However, a fifth item assessing whether or not the participant was an "Appropriate Communicator" remained independent.
Although the full CSRS includes items rating 25 specific conversation behaviors, we had RAs complete a smaller subset of 9 items to provide minimal disruption to the speed-dating process [9] and increase interrater reliability. Of these items, four specific behaviors were deemed most likely to be captured by videos, smiling, humor, vocal confidence, and use of eye contact (Table I).

Finally, the RAs completed 7 items describing their general positive feelings towards the participant (e.g., "My interaction partner and I had a real connection;" "I thought this interaction went smoothly"). As these items were all highly intercorrelated (Cronbach's α=0.927), they were averaged into a global Likeability rating for the interaction.

All ratings were on a 6-point Likert scale with higher scores indicating more effective/positive scores. However, each score only represented a particular RA's reaction to a given participant's behavior. This, in turn, might be shaped by the content of the conversation, the RA's own personal preferences for conversational style, and the RA's unique liking of the participant. As the primary aim of this study was to study broadly effective social skills (i.e., behaviors likely to make a positive impression on multiple people rather than reflecting the preference of a single person), the participant's ratings were averaged across all the RAs they met with in a given round. The mean and standard deviation of the ratings items and the distributions are shown in the Table I. The highest variability among the participants occurs in the Smiling and Humor ratings, whereas the overall appropriateness ratings showed the least variability.

B. *Extracted Features*

We used the SHORE [19] and OpenFace [20] automated facial analysis tools to extract features from each of the interaction videos. The extracted features, listed in Table II, include facial action units (AUs), head pose (position and rotation), and composite expression-based features including angry, happy, sad, surprise, and mouthOpen. These features were extracted for each frame of the video. Each AU is expressed in two forms, a numerical strength value ("_r") and binary presence ("_c").

To analyze the specific importance of when an expression occurs in the interaction, we calculated the averages and variances of each feature across the three disjoint time segments: the first minute of the interaction ("FST"), the middle two minutes of the interaction ("MID"), and the last minute of the interaction ("LST"), as well as over the whole video ("WHL").

V. STATISTICAL ANALYSIS

The original study which gathered the speed-dating data reported that the virtual agent intervention was more successful overall in improving conversational skills ratings than the video intervention [3].

TABLE II. EXTRACTED FEATURES

Feature Extraction Tool	Features
SHORE	happy, angry, sad, surprised, mouthOpen
OpenFace	POSE: head position (Tx,Ty,Tz) head rotation (Rx,Ry,Rz) AUs: 1,2,4,5,6,7,9,10,12,14,15,17,20,23,25,26,45

TABLE III. STRONGEST CORRELATIONS BETWEEN RA RATINGS AND ISOLATED VIDEO FEATURES.

Rating	**Highest Correlated Feature**	***r***	***p-value***
Overall Effectiveness	AU25_r_var in FST	0.498	0.00028
Vocal	AU25_r_var in FST	0.452	0.0034
Smiling	AU06_r_var in LST	0.504	0.00020
Humor	AU26_c_var in LST	0.542	0.00002
Likeability	AU26_c_var in WHL	0.440	0.00600
Appropriate	AU09_c_avg in WHL	-0.467	0.00162
EyeCon	poseRx_var in WHL	-0.360	0.002

In this section, we examine general patterns of associations between these ratings and the videos affective features in a single round.We next analyzed the correlations between feature averages and variances of within each time segment with ratings. As above, all RA ratings and video ratings were averaged across all interactions within a single round, producing 94 unique ratings (first and second round for each of the 47 participants). For each rating, the feature with which it most strongly correlates (Pearson correlation *r*) is displayed in TABLE III. Each rating had a feature which correlated with a magnitude of at least 0.360 (which represents a medium correlation [21]) and ranged as high as 0.542 (which represents a large correlation [21]). Variances of given features (representing dynamic movement of an action unit) tended to correlate more highly than the averages (representing extent of use of that action unit). Both first-minute and last-minute time segments were equally represented.

Variance in lip parting (AU25) over the first minute showed strong correlations with overall effectiveness and vocal confidence ratings, likely highlighting the importance of actively talking right from the beginning of the conversation in making those evaluations. Meanwhile, the smiling and laughing ratings can be captured by variability in the cheek raiser (AU06) and jaw-dropping (AU26) over the last minute, suggesting that the specific evaluation of those behaviors was based on the conversation ending with a variety of genuine smiles and laughter. In contrast, a participant was seen as more likable in a given round when they showed more jaw-dropping (AU26) across the whole interaction.

While most scores were most strongly associated with variability in expressions, it should be noted that RAs were less likely to rate a participant as an Appropriate Communicator in a given round simply based on the average proportion of time spent nose wrinkling (AU09). As this action unit is often used to identify expressions of disgust, it might suggest a more persistent behavior to address.

Although OpenFace does provide a prediction for eye gaze relative to the camera, we were unable to use this feature as the camera positions did not always provide clear views of participants' eyes. Accordingly, we were only able to modestly predict Eye Contact scores using

TABLE IV. P-VALUES OF T-TESTS

	SHORE			
	FST	MID	LST	WHL
FST	-	0.059	0.076	0.388
LST	0.076	0.0004	-	0.034
	OpenFace			
	FST	MID	LST	WHL
FST	-	0.419	0.422	0.024
LST	0.422	0.323	-	0.002

TABLE V. MEAN OF RMSE/SD OVER THE RATING ITEMS

	SHORE	OpenFace
FST	0.982	1.008
MID	0.998	1.026
LST	0.972	0.998
WHL	1.003	1.012

variability of head position (specifically, turning one's head left and right; pose_Rx).

VI. COMPARING TIME SEGMENTS

We next applied LASSO regression analysis [22] (i.e., L1 regularization) with cross-validation to predict each of the conversational skills ratings from the features. The input features were first standardized since the average and variance features were on different scales. We used a dev set to find the optimal L1 regularization parameter (alpha) and using 5-fold cross-validation to predict test-set performance. The analysis was first run separately on SHORE features, and then on OpenFace features for all time segments independently. Shown in Figure 2a and 2b are the root mean squared errors divided by the standard deviation (RMSE/SD) across the test sets using SHORE and OpenFace features respectively. Because each rating has a different variance, by plotting RMSE/SD we are able to represent the accuracy of the prediction independently of the differing natural variation that different rating types entail. Performance in terms of RMSE/SD varied from 0.9 to 1.1. As seen in Figures 2a and 2b, the model performed best at predicting RA ratings of Smiling and Likeability and the worst when predicting Eye Contact ratings. This is expected since several of the features are directly related to smile level (AU6, AU12, AU14, Happy), while none of the features we used directly measured eye gaze. As shown in Figures 2a and 2b regarding likeability, we see that using SHORE features resulted in a lower RMSE/SD that when using OpenFace features . Additionally, using the SHORE features gave better performance in predicting the Smiling and Overall ratings.

Performance varied depending upon which time segment features were used from. When using features from the FST or the LST segment, RMSE/SD was more consistently lower than when using features from either the MID or WHL segments as shown in Figure 2a and 2b. Table V shows the average RMSE/SD for FST, LST, MID, and WHL over the seven ratings items listed in Table I. For both SHORE and OpenFace features LST had the lowest error. We performed unpaired t-test on the errors from different time segments. The p-values are shown in Table IV. Better performance from FST and LST features is consistent with our findings regarding the strongest correlations in Table III which involved FST or LST features for 4 out of 7 of the ratings (see Table IV). These results suggest that evaluations of these skills depend more on information from the first and last minute of the conversation than overall performance.

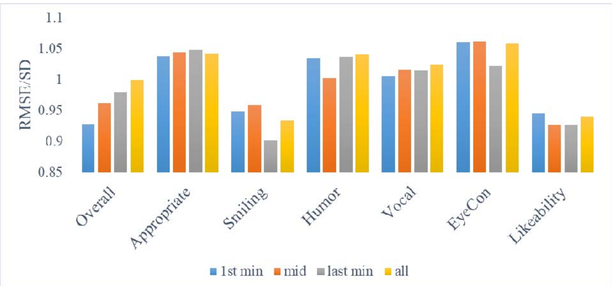
To further evaluate the predictive power of the first and last time segment of the conversation, we then applied LASSO regression analysis using all predictors from both SHORE and OpenFace to compare all indicators in their ability to predict Overall effectiveness ratings. By running the model twice, once with predictors from the first time segment (Figure 2c) and again with predictors from the last time segment (Figure 2d), we were able to evaluate differences in the regression weights that contributed to each prediction.

The most effective predictors of effectiveness in the first time segment was variance in lips parting (AU25) and jaw-dropping (AU26), with more modest contributions by average presence of the dimpler behavior (AU14) and intensity of lips parting (AU25). Taken together, these behaviors suggest that overall effectiveness is predicted by talking in the first minute without any consideration for emotional valence. In comparison, Figure 2d highlights a variety of affective considerations including variability in jaw-dropping (AU26), happy expressions, and avoidance of nose wrinkling (AU09). As these action units were related to laughing and disgust expressions, this suggests that it is important to end a conversation with effusive expressions of warmth. An identical analysis was repeated with the support vector machine (SVM) model with a radial basis function kernel and cross-validated hyperparameters. We found that SVM was consistently outperformed by LASSO. This is likely due to the limited size of our dataset. The SVM model with RBF kernel is a more complex model which needs more data than LASSO in order to train effectively. A fundamental element of LASSO is that it will minimize the number of input dimensions used, effectively encouraging simpler models. Thus, it is likely LASSO is providing us better results since it is less likely to overfit the limited data.

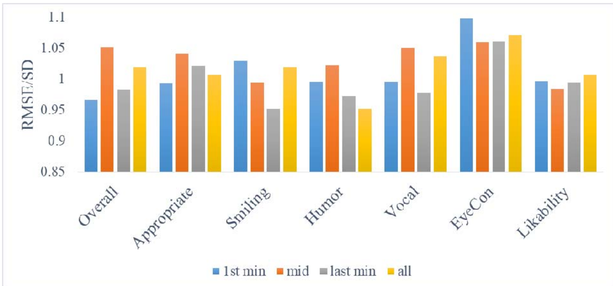
VII. DISCUSSION

The present results highlight that – when evaluating specific nonverbal behaviors – timing matters. In particular, the facial expressions used in the first and last minute of a conversation have the most important contributions to perceived effectiveness. These findings are consistent with prior research and are suggestive of classic primacy and recency effects [23], underscoring the need for feedback mechanisms to provide dynamic feedback based on context. The stark contrasts in behaviors highlighted by LASSO regression in the first and last minute (Figure 2c and 2d) delineate the different behaviors that might be most critical at the beginning and end of such conversations. In particular, LASSO regressions highlighted the key role of immediately speaking and establishing vocal confidence in the first minute However, as a conversation comes to its close, the strongest predictor of overall effectiveness were expressions associated with laughter, happiness, and the absence of disgust expressions. These suggest the emotional dimensions of a discussion matter most in its last minute.

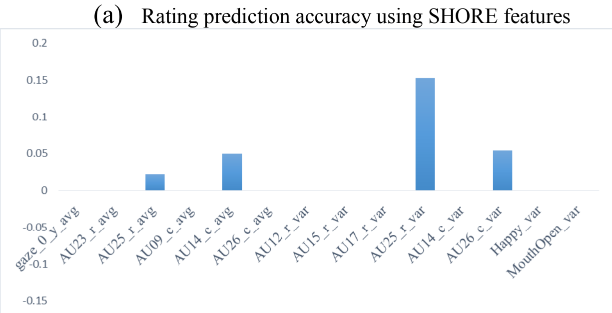
These findings highlight that both primacy and recency effects seem to have a role in determining the RA evaluations of effectiveness, but that different behaviors are valued at different times.



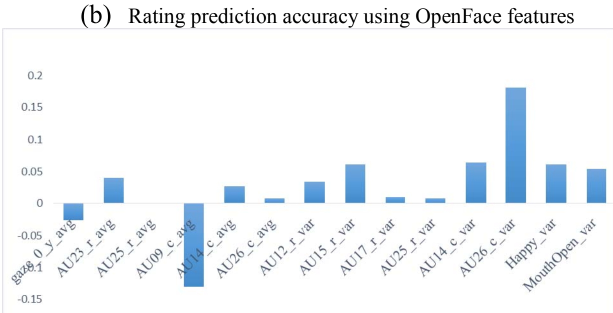
(a) Rating prediction accuracy using SHORE features



(b) Rating prediction accuracy using OpenFace features



(c) LASSO linear regression feature weights from FST time segment



(d) LASSO linear regression features weights from LST time segment

Figure 2. LASSO linear regression performance and feature weights

To understand the role of these behaviors, it is helpful to use the framework of "competence" and "warmth" from social cognitive research on person perception [24]. The theory holds that evaluations of "warmth" – people's kindness and good intentions – lead to liking while evaluations of "competence" – people's ability carry out their aims – leads to respect. It seems that a speaker's goals in conversation should shift along these dimensions so that they establish competence early on and then build a friendship with expressions of warmth.

While promising, these results are bounded by some limitations. The video ratings were only gathered from male college students in a highly artificial context (i.e., time-bound "speed-dates" with generally encouraging research staff). While this afforded us the precision for the current analyses, future studies should investigate generalizability of these findings across a diverse group of individuals in more naturalistic, unbounded settings, where the beginning and ending of a conversation do not occur in fixed timeframes. Additionally, the current prediction models are built entirely from affective visual features. It is likely that richer prediction can be achieved by adding robust gaze tracking and voice analysis features. Finally, the current approach simply views the target participant in isolation without considering the interaction partner. Given the interpersonal nature of social effectiveness, a more dynamic understanding of what makes effectiveness work can be obtained by examining the correspondence of features across the dyad [25].

VIII. CONCLUSION

We present an automated prediction framework for automatically rating several types of conversational skills from visual affective features. The models used show encouraging results for predicting many of these ratings and demonstrate that not only was a prediction from a time segment possible, but it resulted in better prediction than evaluating the whole video when using the first or last segment of the interaction. Additionally, as different features show different levels of importance across separate time segments of the conversation. These results provide the framework for the design of dynamic nonverbal feedback agents that provide differing feedback based on differing goals at different points in a conversation.

ACKNOWLEDGEMENT

The work was supported by NSF EAGER grant IIS-1543758

REFERENCES

[1] D. R. Carney, C. R. Colvin, and J. A. Hall, "A thin slice perspective on the accuracy of first impressions," J. Res. Pers., vol. 41, no. 5, pp. 1054–1072, 2007.

[2] Alan C. Fox, "People Tools: 54 Strategies for Building Relationships, Creating Joy, and Embracing Prosperity", 2014.

[3] M.R. Ali, D. Crasta, L. Jin, A. Baretto, J. Pachter, R.D. Rogge and M.E. Hoque, "LISSA Live Interactive Social Skills Assistance" Proc. IEEE Affective Computing and Intelligent Interaction, pp. 173–179, 2015.

[4] B.H. Spitzberg, "Conversational skills rating scale: An instructional assessment of interpersonal competence." Annandale, VA: Speech Communication Assoc. 1995.

[5] D. Gatica-Perez, "Automatic nonverbal analysis of social interaction in small groups: A review," Image Vis. Comput., vol. 27, no. 12, pp. 1775–1787, 2009.

[6] L. Batrinca, G. Stratou, A. Shapiro, L. P. Morency, and S. Scherer, "Cicero - Towards a multimodal virtual audience platform for public speaking training," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8108 LNAI, pp. 116–128, 2013.

[7] T. Baur, I. Damian, F. Lingenfelser, J. Wagner, and E. André, "NovA: Automated analysis of nonverbal signals in social interactions," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8212 LNCS, pp. 160–171, 2013.

[8] M.E. Hoque, M. Courgeon, J.-C. Martin, B. Mutlu, and R.W. Picard, "Mach: My automated conversation coach," Proc. 2013 ACM Int. Jt. Conf. Pervasive ubiquitous Comput. - UbiComp '13, p. 697, 2013.

[9] Finkel, E. J., Eastwick, P. W., & Matthews, J. (2007). Speed‐dating as an invaluable tool for studying romantic attraction: A methodological primer. Personal Relationships, 14(1), 149-166.

[10] E.J. Finkel and P.W. Eastwick, "Speed-dating." Curr. Directions in Psych. Sci,, 17(3), 193-197, 2008.

[11] A. Veenstra and H. Hung, "Do they like me? Using video cues to predict desires during speed-dates," Proc. IEEE Int. Conf. Comput. Vis., pp. 838–845, 2011.

[12] R. Ranganath, D. Jurafsky, and D. A. McFarland, "Detecting friendly, flirtatious, awkward, and assertive speech in speed-dates," Comput. Speech Lang., vol. 27, no. 1, pp. 89–115, 2013.

[13] R. Ranganath, D. Jurafsky, and D. Mcfarland, "It's Not You, it's Me: Detecting Flirting and its Misperception in Speed-Dates," EMNLP '09 Proc. 2009 Conf. Empir. Methods Nat. Lang. Process., vol. 1, no. August, pp. 334–342, 2007.

[14] T. Vacharkulksemsuk, E. Reit, P. Khambatta, P. W. Eastwick, E. J. Finkel, and D. R. Carney, "Dominant, open nonverbal displays are attractive at zero-acquaintance," Proc. Natl. Acad. Sci., vol. 201508932, no. 15, pp. 1–6, 2016.

[15] M. Weisbuch, Z. Ivcevic, and N. Ambady, "On being liked on the web and in the "real world": Consistency in first impressions across personal webpages and spontaneous behavior", In Journal of

Experimental Social Psychology, Volume 45, Issue 3, 2009, Pages 573-576.

[16] J. R. Curhan and A. Pentland, “Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes.,” J. Appl. Psychol., vol. 92, no. 3, pp. 802–811, 2007.

[17] “5 Body Language Tricks To Make Anyone Instantly Like You - Personality Development & English Lessons.” .

[18] D. Wendler, Improve your social skills. 2014.

[19] A. Ernst, T. Ruf, and C. Kueblbeck, “A Modular Framework to Detect and Analyze Faces for Audience Measurement Systems,” Proc. Work. Perad., pp. 75–87, 2009.

[20] T. Baltruˇ and P. Robinson, “OpenFace: an open source facial behavior analysis toolkit.”

[21] S. Sawilowsky, "New effect size rules of thumb.". Journal of Modern Applied Statistical Methods. 8 (2): 467–474 2009.

[22] R. Tibshirani, “Regression Selection and Shrinkage via the Lasso,” Journal of the Royal Statistical Society B, vol. 58, no. 1. pp. 267–288, 1996.

[23] S. Labbie, “Primacy-recency effects in impression formation and congruity-incongruity of stimulus material,” Percept. Mot. Skills, vol. 37, no. 1, pp. 275–278, 1973.

[24] S. T. Fiske, A. J. C. Cuddy, and P. Glick, “Universal dimensions of social cognition: warmth and competence,” Trends Cogn. Sci., vol. 11, no. 2, pp. 77–83, 2007.

[25] L. M. van Swol, “The effects of nonverbal mirroring on perceived persuasiveness, agreement with an imitator, and reciprocity in a group discussion,” Communic. Res., vol. 30, no. 4, pp. 461–480, 2003.