

Predicting Client’s inclination towards Target Behavior Change in Motivational Interviewing and investigating the role of laughter

Rahul Gupta¹, Panayiotis G. Georgiou¹, David C. Atkins² and Shrikanth Narayanan¹

¹Signal Analysis and Interpretation Lab (SAIL), Department of Electrical Engineering
University of Southern California, Los Angeles, CA

²Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA
guptarah@usc.edu, georgiou@sipi.usc.edu, datkins@u.washington.edu, shri@sipi.usc.edu

Abstract

Motivational interviewing (MI) is a goal oriented psychotherapy that facilitates intrinsic motivation within a client in order to change behavior in a dialog setting. The Motivational Interviewing Skills Code (MISC) is a manual observational coding method used to quantify and evaluate the quality of MI sessions using their audio-visual recordings. However, this coding method is both labor intensive and expensive. We present an approach towards automating MISC assignments in MI involving addiction cure. Specifically, we focus on predicting valence for “Client Change Talk” (ChangeTalk) utterances, which indicate a client’s attitude towards a “Target Behavior Change” (Target). We further study the effect of incorporating counselor behavior in the model. We observe that our best model achieves an unweighted accuracy of 50.8% in a 3-way classification of positive vs negative valence ChangeTalk vs no ChangeTalk. Furthermore, we study the effect of including non-verbal behavior, specifically laughters, in our model. Information regarding location of laughters improves the unweighted accuracy of our model to 51.4% and our experimental results suggest prosodic differences in laughters belonging to ChangeTalk utterances with different valences.

Index Terms: Motivational Interviewing (MI), Motivational Interviewing Skill Code (MISC), Client Change Talk (ChangeTalk) utterance, Maximum entropy model.

1. Introduction

Rollnick et al. [1] define Motivational Interviewing (MI) as “a directive, client-centered counseling style for eliciting behavior change by helping clients to explore and resolve ambivalence”. MI setting is extensively used in addiction related problems (alcohol, drugs, etc.) [2, 3]. MI helps addiction patients perceive both, the benefit (e.g., the high) and the harm (e.g., health, personal problems) and helps them resolve the ambivalence in a dyadic spoken interaction with a therapist towards a positive change. With emerging evidence base and popularity for MI [4], a significant challenge lies in ensuring high quality of treatment which calls for standard metric to assess quality of such interactions. The Motivational Interviewing Skill Code (MISC) [5] has emerged as a standard observational method for evaluating the quality of MI interactions. MISC is a behavioral coding systems in which human coders are trained to annotate video or audio tapes over several parameters as global counselor ratings, empathy, behavior categories etc.. The uses of MISC range from providing detailed session feedback to counselors in the process of learning MI to predicting treatment outcome from psychotherapy measures.

MISC provides us with a systematized tool to assess the quality of MI interactions. However such manual coding systems are not scalable to real world use due to time, labor and economic constraints [6]. As part of our ongoing efforts towards Behavioral Signal Processing [7, 8], in this paper, we present a model towards the automation of MISC annotation. Specifically, we focus on the “Target Behavior Change” (Target) aspect of the client behavior that specifies a target client behavior (smoking, drinking, medication) and a direction of change (stopping, increasing etc.). We conduct three sets of experiments to (a) predict a client’s attitude towards a Target, (b) investigate the effect of laughters during interaction on Target and (c) look for prosodic differences in laughters with respect to different behaviors.

(a) *Predicting client’s attitude towards Target:* A turn in client speech is annotated with a positive/negative valence if it shows an inclination towards/away from Target. Such client turns are termed as “Client Change Talk” (ChangeTalk) utterances. We design a lexical based model for predicting positive vs negative valence ChangeTalk vs no ChangeTalk given a client utterance. Additionally, we study the effect of including counselor behavior in our prediction model. Each counselor utterance is MISC annotated with a behavior code (reflect, support etc.). As the counselors are required to carefully attend to client language related to the target behavior, their behavior may carry indicators to the client’s attitude towards Target. Our best prediction model achieves an unweighted accuracy (UA) of 50.8% (chance 33.3%) for the three way classification.

(b) *Investigating the role of laughters during interaction:* Laughters are linked to human behavior and are hypothesized to carry out several social functions [9, 10]. In this experiment, we examine if the mere occurrence of a laughter during interaction carries some information with regards to the ChangeTalk valence. We observe that inclusion of information regarding their presence improves UA of our previous model to 51.4%.

(c) *Prosodic differences amongst laughters:* Several studies [11, 12, 13] suggest differences amongst laughters contingent upon the context in which they happen. We look for prosodic differences amongst client laughters belonging to utterances in the three ChangeTalk classes. We observe some discriminatory power in their prosody and achieve an UA of 40.5% in classifying laughters belonging to utterances from the three classes.

2. Database

Our experiments pertain to an MI based intervention study on drug abuse problems involving patients at a public, safety-net hospital [14, 21, 22]. We use data from 49 subjects coded

Utterance	Counselor behavior code	ChangeTalk
T: Did you come here on own? C: Yes, I was sure about this.	Question	Change+
T: I am here to help you. C: I appreciate that.	Support	Change0
C: I really have to stop drugs, C: but I just don't want to.		Change+ Change-

Table 1: Example conversation with Counselor (T) behavior annotation and client (C) ChangeTalk valence annotation.

Counselor behavior codes	
Advice (AD) (101)	Raise Concern (RC) (3)
Affirm (AF) (427)	Filler (FI) (29)
Confront (CF) (4)	Reflect (RE) (2310)
Direct (DI) (2)	Reframe (RF) (0)
Emphasize control (EC) (36)	Support (SU) (260)
Facilitate (FA) (4175)	Structure (ST) (215)
Giving information (GI) (1475)	Warn (WA) (5)
Question (QU) (1665)	
Client change talk utterance	
Positive valence (1749)	Negative valence (1253)
No ChangeTalk in utterance (9060)	

Table 2: List of various counselor behavior codes and ChangeTalk codes and corresponding count of utterances.

by a single annotator as per the MISC manual. We present an excerpt from one of the sessions and the corresponding ChangeTalk and counselor behavior codes in Table 1. Utterances with positive valence during ChangeTalk are listed as Change+, negative valence as Change- and with no change talk as Change0. Table 2 lists the count of utterances belonging to various ChangeTalk and counselor behavior codes.

3. Experiments

3.1. Predicting valence of Client Change Talk utterance.

Although modeling the MISC annotation of an MI session may be very complex, in this work we propose a simplified scheme to address this problem. We represent each counselor utterance during the interaction using the set of variables $\{U_{T1}, \dots, U_{TN}\}$ and each client utterance as $\{U_{C1}, \dots, U_{CM}\}$. These utterances are the observed outcomes determined by complex and time varying internal mental processes of each participant involving a mix of several affective, attitudinal, emotional states etc. Block (a) in Figure 1 lays out this representation of the dyadic interaction. During annotation, the coder observes each of the client and counselor utterances and assigns a code. The codes that we focus on are intended to capture specific behavior without regards to the overall impression of MI. Coders are made aware of the target behavior in detail so that they can reliably discriminate it from all other topics. In the following sections, we describe our baseline methodology to predict ChangeTalk valences and the additions we make to evaluate the effect of including counselor behavior codes.

3.1.1. Baseline system

Since the coder assigns the ChangeTalk codes after individually observing each of the client utterances, a simplistic model may assume that the assigned code is determined solely by the utterance in question. We design our baseline system based on

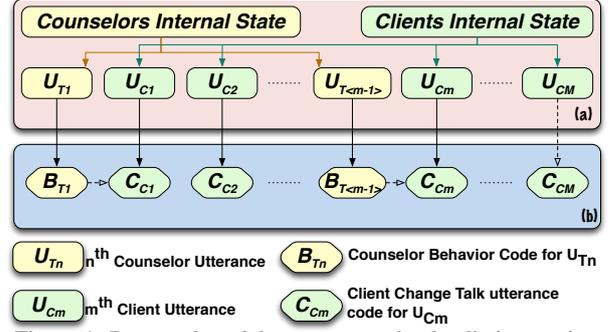


Figure 1: Proposed model to represent the dyadic interaction and the annotation process.

this assumption. In this baseline model, we focus on lexical content of client utterances and design an n-gram based classification system. We describe our feature extraction, selection and classification scheme below. We perform a leave one session out cross validation in all our experiments.

(1) *Feature extraction*: Given the unbalanced class distribution, we initially downsample instances from the majority classes (Change0, Change+), so that each class has the same number of instances as the least represented class; Change-. We extract all the unigrams and bigrams from each client utterances as potential features to learn the corresponding ChangeTalk code. However, this leads to a large feature space with several features that may not be relevant to the classification or may contain n-grams that are rare in occurrence. To overcome this problem, we use a feature selection algorithm as described next.

(2) *Feature selection*: We select a given n-gram (ng_k) based on the entropy ($E(ng_k)$, equation 1) of its empirical distribution (equation 2) over the three classes and their minimum count ($\#ng_k$) on the downsampled data as shown in the Algorithm 1. The maximum entropy threshold (T_E) and the minimum word count (C_{\min}) thresholds in the algorithm are tuned on a subset of the training set.

Algorithm 1 N-gram selection for classifier training.

- 1: Define: $G = \{ng_1, \dots, ng_k, \dots, ng_K\}$: Set of n-grams in the training set
- 2: G_{sel} : Set of selected n-grams for classification
- 3: Initialize: $G_{\text{sel}} = \Phi$
- 4: **for** $k = 1 \dots K$ **do**
- 5: **if** ($E(ng_k) < T_E$) and ($\#ng_k > C_{\min}$) **then**
- 6: $G_{\text{sel}} = G_{\text{sel}} \cup ng_k$
- 7: **end if**
- 8: **end for**

$$E(ng_k) = \sum_{\substack{\text{ChangeTalk}_{\text{val}} \in \\ \{\text{Change+}, \text{Change-}, \text{Change0}\}}} P_{k, \text{ChangeTalk}_{\text{val}}} \log(P_{k, \text{ChangeTalk}_{\text{val}}}) \quad (1)$$

$$\text{where, } P_{k, \text{ChangeTalk}_{\text{val}}} = P(ng_k / \text{ChangeTalk}_{\text{val}}) = \frac{\#ng_k \in \text{Training utterances coded as ChangeTalk}_{\text{val}}}{\#ng_k \in \text{Training set}} \quad (2)$$

(3) *Classifier training*: We train a conditional maximum entropy model [15] on the chosen set of n-grams (G_{sel}). We perform the parameter estimation using L-BFGS [16]. Our baseline system can be viewed as only utilizing the solid line connecting U_{Cm} to C_{Cm} in Figure 1, i.e. ChangeTalk code depends only on what client says. Equation 3 shows the decision

rule for assigning class the **ChangeTalk** class C_{Cm} to the utterance U_{Cm} given a set of observed features O_{Cm} . In our baseline model, we set O_{Cm} to $G_{sel,Cm}(\subset G_{sel})$, the set of n-grams extracted from the considered utterance U_{Cm} .

$$C_{Cm} = \arg \max_{\text{ChangeTalk}_{val}} P(\text{ChangeTalk}_{val}/O_{Cm}) \quad (3)$$

$\in \{\text{Change+}, \text{Change-}, \text{Change0}\}$

3.1.2. Classification system incorporating counselor behavior

In our next experiment, we account for the counselor behavioral codes while inferring the **ChangeTalk** codes, i.e. the client’s utterance and the counselor’s immediate past behavior both contribute towards identifying the code. Hence, in this scheme we also utilize the dotted link connecting the preceding counselor behavior code $B_{T<m-1>}$ in addition to features from U_{Cm} to infer C_{Cm} . We perform two sets of experiments incorporating oracle and inferred counselor behavior codes as follows.

Oracle counselor behavior: In this experiment, we use the oracle counselor behavior code ($B_{T<m-1>}^{orc}$) preceding C_{Cm} as annotated by the coder. This model utilizes $\{G_{sel,Cm}, B_{T<m-1>}^{orc}\}$ as the evidence O_{Cm} in equation 3. Note this is not ideally possible in a real system as we are using the true counselor codes for the test set which likely won’t be available in the real world scenario.

Inferred counselor behavior: As the use of oracle values for oracle counselor behavior is impractical, we develop a system to infer them. We implement the same framework as described in the baseline system. However as we have extremely low number of training instances for a few codes, we merge several minority classes before training our prediction system. We retain questions (*QU*), giving information (*GI*), reflection (*RE*) and facilitate (*FA*) and merge all the other classes into a fifth class; others (*OT*). Initially, we gauge the effect of merging counselor codes on our previous prediction system involving oracle counselor codes. O_{Cm} in equation 3 is set to $\{G_{sel,Cm}, B_{T<m-1>}^{orc,mer}\}$, where $B_{T<m-1>}^{orc,mer}$ are the oracle counselor codes obtained after merging.

In order to predict the merged counselor codes, we initially downsample the data to remove class bias over the 5 classes. We perform feature selection and classifier training as described in the baseline system for inferring the counselor codes using the n-grams from counselor utterances (U_{Tn}). Equation 4 shows the rule for inferring the counselor code $B_{T<m-1>}^{inf,mer}$ from the set of features $O_{T<m-1>}$. In this model, $O_{T<m-1>}$ is set to the selected set of n-grams $G_{sel,T<m-1>}$ extracted from counselor utterance $U_{T<m-1>}$ preceding U_{Cm} . We use this inferred counselor code as the observed evidence O_{Cm} in equation 3 ($O_{Cm} = \{G_{sel,Cm}, B_{T<m-1>}^{inf,mer}\}$).

$$B_{T<m-1>}^{inf,mer} = \arg \max_{B_{val} \in \{RE, GI, QU, FA, OT\}} P(B_{val}/G_{sel,T<m-1>}) \quad (4)$$

3.1.3. Results and discussion

We use unweighted accuracy (UA) as our evaluation metric and also report the F-measure for Change+ and Change- given their low occurrence frequency relative to Change0. The results for inferring **ChangeTalk** codes using the baseline system and after incorporating $B_{T<m-1>}$, $B_{T<m-1>}^{orc,mer}$ and $B_{T<m-1>}^{inf,mer}$ are shown in Table 3. We also show the results for inferring counselor behavior codes in Table 4.

Discussion: From the results we observe that our model performs well above chance, indicating that just lexical content during conversation can inform us of change talk behavior. Improvement in the results, after including the counselor behavior

Model	UA	Change+		Change-	
		F-meas.	Accuracy/Precision	F-meas.	Accuracy/Precision
Chance	33.3	20.2	33.3/14.5	15.8	33.3/10.4
U_{Cm}	49.0	30.3	44.8/22.9	28.5	49.2/20.1
$+B_{T<m-1>}^{orc}$	50.8	32.7	44.8/25.8	29.2	49.1/20.8
$+B_{T<m-1>}^{orc,mer}$	50.6	32.6	44.7/25.6	29.1	48.8/20.7
$+B_{T<m-1>}^{inf,mer}$	50.2	31.8	43.6/25.0	29.1	48.5/20.8

Table 3: Results (in %) for predicting **ChangeTalk** codes.

Model	UA	Class Accuracies				
		RE	GI	QU	FA	OT
Chance	20	20	20	20	20	20
U_{Tn}	69.6	70.7	55.7	77.2	97.1	47.2

Table 4: Results (in %) for predicting counselor behavior codes.

code indicates, that coder does take context of conversation into account while assigning the **ChangeTalk** codes. Also, we do not observe any significant difference between using all the behavior codes versus using merged oracle codes. This suggests that training the model on a few codes with sufficient number of samples is as good as training on all the codes with a few samples. However, we do observe a decrease when we use the inferred counselor codes over oracle codes. This stems from the imperfect prediction of the counselor codes themselves. We observe that a few classes in case of counselor codes are predicted more accurately as compared to others in spite of using balanced number of instances during training. This indicates that some classes are better distinguishable with lexical features, while other classes may be more distinguishable in other modalities. Particularly in the case of facilitate (*FA*), as most of utterances are simple, functioning as “keep going” acknowledgment such as “Mm Hmm”, “OK” etc., we observe almost perfect prediction.

3.2. Prediction incorporating laughters

Several studies suggest the importance of laughters in discourse [9, 10]. In this section, we perform preliminary analysis of laughters and study their effect on our previous system. We hypothesize that mere occurrence of laughter events may provide us with some information regarding a client’s attitude regarding **Target**. We add a simple binary feature indicating presence of laughter L_{Cm} (available through transcripts) in the client utterance U_{Cm} to our previous models and reproduce the results. We set O_m to $\{G_{sel,Cm}, L_{Cm}\}$ for the baseline model and the similar addition is made to models incorporating the counselor behavior. While inferring $B_{T<m-1>}^{inf,mer}$, we use a binary features $L_{T<m-1>}$ indicating the presence of the counselor laughter in $U_{T<m-1>}$ ($O_{T<m-1>} = \{G_{sel,T<m-1>}, L_{T<m-1>}\}$ in equation 4). We show the results for predicting **ChangeTalk** codes using laughters and relative improvements over the counterparts from the previous model in Table 5. Results and corresponding relative improvements for counselor behavior code prediction are shown in Table 6.

Discussion: From the results, we observe that we get a consistent gain in both the client and the counselor results. We list the empirical distribution of client laughters over the three **ChangeTalk** codes in Figure 2(a) and the counselor laughters over the merged counselor behavior codes in Figure 2(b). The occurrence probability of laughters is not uniform over codes, thereby providing additional information to the previous model. We observe that in the case of **ChangeTalk** codes, an utterance labeled Change+ is most likely to contain laughters. This indi-

Model	UA	Change+		Change-	
		F-meas.	Acc./Prec.	F-meas.	Acc./Prec.
U_{Cm}	50.1 (2.2)	32.1 (5.9)	45.8/24.7 (2.2/7.7)	28.6 (0.0)	49.2/20.1 (0.0/0.0)
$+B_{T<m-1}^{orc}$	51.4 (1.2)	33.4 (2.1)	45.4/26.4 (1.3/2.3)	29.6 (1.7)	49.5/21.1 (0.8/1.4)
$+B_{T<m-1}^{orc,mer}$	51.3 (1.4)	33.4 (2.5)	45.3/26.4 (1.3/3.1)	29.5 (1.4)	49.5/21.0 (1.4/1.4)
$+B_{T<m-1}^{inf,mer}$	50.8 (1.2)	32.8 (3.1)	44.1/26.1 (1.1/4.4)	29.3 (0.6)	48.9/20.9 (0.8/0.4)

Table 5: Results (%) & relative improvements (%) over previous model for predicting ChangeTalk codes w/ laughter.

Model	UA	Class Accuracies				
		RE	GI	QU	FA	OT
U_{Tn}	71.3 (2.4)	71.4 (1.0)	56.0 (0.5)	77.2 (0.0)	97.0 (-0.1)	55.1 (16.7)

Table 6: Results (%) & relative improvements (%) over previous model for predicting counselor behavior codes w/ laughter.

cates that presence of laughters shows a positive attitude. The occurrence probability of laughters is more skewed in the case of counselor codes. We observe that the utterance labeled *OT* are most likely to contain laughter. Due to this, fact we observe largest increase in the classification accuracy for the *OT* class. Laughters are rarely present in utterances labeled *FA* and hence we barely see any change in its class accuracy.

3.3. Laughters and their prosodic differences

Several studies show that there are inherent differences in laughters and the context in which they occurs [11, 12, 13]. We investigate the differences amongst client laughters in the context of the three ChangeTalk classes. We design a prosody based classification system to distinguish amongst laughters that occur in utterances coded as Change+, Change- or Change0. We hypothesize that if indeed differences exist amongst the laughters, this can further aid our ChangeTalk valence prediction system. We describe the prosodic features and the classification system below.

Features: We manually annotate all the 371 client laughters (Change+:61, Change-:26, Change0: 284) in the 49 sessions marking their start and end positions. Given a small number of samples, we limit our experiment to a few low level prosodic cues and compute their global statistics shown in Table 7. We mean normalize these features per speaker.

Classifier: We use a linear kernel support vector machine classifier. Given the unbalanced class distribution (counts shown in Figure 2), we downsample the samples from Change0 and Change+ classes so that each class has equal number of samples. We perform leave one session out cross-validation on the laughters from 49 sessions. We list the classification accuracy in Table 8.

Discussion: We observe that the use of a few low level prosodic cues does provide us with some discriminatory power in between client laughters belonging to the three classes. Poor classification accuracy stems from extremely small number of training instances. Laughters from the class Change0 are most poorly classified as they have a high downsampling factor for maintaining the class balance. This introduces a sampling bias. Due to the same reason, we could not carry out an experiment to investigate differences in counselor laughters as some classes have too few samples (e.g. 7 for QU). Because of the

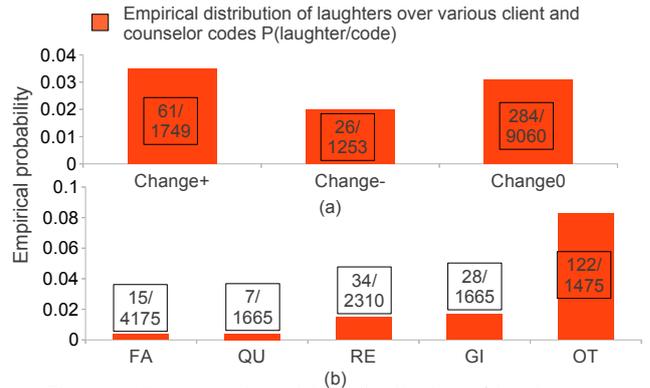


Figure 2: Counts and empirical distribution of laughters over (a) ChangeTalk codes (b) Counselor behavior codes.

Feature	Statisticals
Pitch, intensity, voicing probability, harmonic to noise ratio	Mean, variance, range

Table 7: Prosodic features used in laughter classification.

Model	Unweighted accuracy	Class accuracies		
		Change+	Change-	Change0
Chance	33.3	33.3	33.3	33.3
Prosody	40.5	47.5	47.2	26.9

Table 8: Results for classification of client laughters over ChangeTalk codes using prosody.

weak discrimination, this information does not help our previous ChangeTalk valence prediction model as of now. However, given that prosodic differences in laughters do exist, we are encouraged that we improve the proposed model with the availability of more training data in future.

4. Conclusion

In this paper, we present a scheme towards automatically obtaining MISC codes in MI settings. We design a lexical based scheme to automatically identify client utterances with positive or negative valence that indicate their attitude towards a targeted behavior change. We show that incorporating the counselor's behavior into account during the interaction helps improve the prediction, thus validating the importance of the counselor towards positive outcomes. We proceed with a preliminary analysis on incorporating laughters as additional information source, augmenting the previous system. Finally, we analyze the type of laughters based on their prosodic cues. We observe some discriminatory power in the prosody of laughters with respect to the ChangeTalk valences, however due to limited data and poor accuracy we could not exploit this towards change talk classification.

We presented our results on one aspect of MISC code. However, the MISC annotations provide other global measures such as empathy, motivational interviewing spirit etc. that furnish more indicators regarding the success of a session. We aim at building upon our current system to incorporate these measures. Studies link acoustic [17, 18], visual cues [19, 20] etc. to human behavior and one can incorporate such cues to supplement the system. Also, we aim to further investigate other aspects of laughters (e.g. shared laughters) and the role they play in MI. One may further extend this work to other non-verbal vocalizations as fillers, sighs etc..

5. Acknowledgement

This research is supported by NSF, NIH and DARPA.

6. References

- [1] Stephen Rollnick and William R Miller, "What is motivational interviewing?," *Behavioural and cognitive Psychotherapy*, vol. 23, no. 04, pp. 325–334, 1995.
- [2] William R Miller and Stephen Rollnick, *Motivational interviewing: preparing people for change*, New York [etc.]: The Guilford Press, 2002.
- [3] John S Baer, David B Rosengren, Christopher W Dunn, Elizabeth A Wells, Richard L Ogle, and Bryan Hartzler, "An evaluation of workshop training in motivational interviewing for addiction and mental health clinicians," *Drug and alcohol dependence*, vol. 73, no. 1, pp. 99–106, 2004.
- [4] Brian L Burke, Christopher W Dunn, David C Atkins, and Jerry S Phelps, "The emerging evidence base for motivational interviewing: a meta-analytic and qualitative inquiry," *Journal of Cognitive Psychotherapy*, vol. 18, no. 4, pp. 309–322, 2004.
- [5] William R Miller, Theresa B Moyers, Denise Ernst, and Paul Amrhein, "Manual for the motivational interviewing skill code (misc)," *Unpublished manuscript. Albuquerque: Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico*, 2003.
- [6] Enola Proctor, Hiie Silmere, Ramesh Raghavan, Peter Hovmand, Greg Aarons, Alicia Bunger, Richard Griffey, and Melissa Hensley, "Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda," *Administration and Policy in Mental Health and Mental Health Services Research*, vol. 38, no. 2, pp. 65–76, 2011.
- [7] Shrikanth S Narayanan and Panayiotis G Georgiou, "Behavioral signal processing: Deriving human behavioral informatics from speech and language," *Proceedings of the IEEE*, vol. 101, no. 5, pp. 1203–1233, 2013.
- [8] Panayiotis G Georgiou, Matthew P Black, Adam C Lammert, Brian R Baucom, and Shrikanth S Narayanan, "That's aggravating, very aggravating: Is it possible to classify behaviors in couple interactions using automatically derived lexical features?," in *Affective Computing and Intelligent Interaction*, pp. 87–96. Springer, 2011.
- [9] Rose L Coser, "Some social functions of laughter," *Human relations*, 1959.
- [10] Helmuth Plessner, *Laughing and crying: a study of the limits of human behavior*, Northwestern University Press, 1970.
- [11] Viveka Adelswärd, "Laughter and dialogue: The social significance of laughter in institutional discourse," *Nordic Journal of Linguistics*, vol. 12, no. 2, pp. 107–136, 1989.
- [12] Phillip J Glenn, *Laughter in interaction*, Cambridge University Press Cambridge, 2003.
- [13] Theodora Chaspari, Emily M Provost, Athanasios Katsamanis, and Shrikanth S Narayanan, "An acoustic analysis of shared enjoyment in eca interactions of children with autism," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4485–4488.
- [14] Antoinette Krupski, Jutta M Joesch, Chris Dunn, Dennis Donovan, Kristin Bumgardner, Sarah Peregrine Lord, Richard Ries, Peter Roy-Byrne, et al., "Testing the effects of brief intervention in primary care for problem drug use in a randomized controlled trial: rationale, design, and methods," *Addiction science & clinical practice*, vol. 7, no. 1, pp. 27, 2012.
- [15] Zhang Le, "Maximum entropy modeling toolkit for python and c++," *Natural Language Processing Lab, Northeastern University, China*, 2004.
- [16] Dong C Liu and Jorge Nocedal, "On the limited memory bfgs method for large scale optimization," *Mathematical programming*, vol. 45, no. 1-3, pp. 503–528, 1989.
- [17] Daniel Bone, Chi-Chun Lee, Theodora Chaspari, Matthew P Black, Marian E Williams, Sungbok Lee, Pat Levitt, and Shrikanth Narayanan, "Acoustic-prosodic, turn-taking, and language cues in child-psychologist interactions for varying social demand," 2013.
- [18] Matthew P Black, Athanasios Katsamanis, Brian R Baucom, Chi-Chun Lee, Adam C Lammert, Andrew Christensen, Panayiotis G Georgiou, and Shrikanth S Narayanan, "Toward automating a human behavioral coding system for married couples interactions using speech acoustic features," *Speech Communication*, vol. 55, no. 1, pp. 1–21, 2013.
- [19] Alexandra Frischen, Andrew P Bayliss, and Steven P Tipper, "Gaze cueing of attention: visual attention, social cognition, and individual differences," *Psychological bulletin*, vol. 133, no. 4, pp. 694, 2007.
- [20] Bo Xiao, Panayiotis G Georgiou, Brian R Baucom, and Shrikanth S Narayanan, "Power-spectral analysis of head motion signal for behavioral modeling in human interaction," in *Proceedings of IEEE International Conference on Audio, Speech and Signal Processing (ICASSP)*, May 2014.
- [21] Bo Xiao, Panayiotis G Georgiou, and Shrikanth S Narayanan, "Analyzing the language of therapist empathy in motivational interview based psychotherapy," *Proc Asia Pacific Signal Inf Process Assoc*, pp. 1–4, 2012.
- [22] Doğan Can, Panayiotis G Georgiou, David C Atkins, and Shrikanth S Narayanan, "A case study: Detecting counselor reflections in psychotherapy for addictions using linguistic features," in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.