Developing an Automated Report Card for Addiction Counseling: The Counselor Observer Ratings Expert for MI (CORE-MI)

James Gibson
School of Engineering
University of Southern California
Los Angeles CA
jjgibson@usc.edu

Geoff Gray
Division of Design
University of Washington
Seattle WA
gegray@uw

Tad Hirsch
Division of Design
University of Washington
Seattle, WA
thirsch@uw.edu

Zac E. Imel
Department of Educational Psychology
University of Utah
Salt Lake City UT
zeimel@gmail.com

Shrikanth Narayanan
School of Engineering
University of Southern California
Los Angeles CA
shri@sipi.usc.edu

David Atkins
Department of Psychiatry
University of Washington
Seattle WA
datkins@uw.edu

ABSTRACT
We present an overview of CORE-MI, an automated coding and visualization system that provides report-card like feedback on psychotherapists' adherence with Motivational Interviewing (MI). MI is a widely studied and effective method of substance abuse treatment, and is widely practiced. However, there are currently no technologies to provide immediate, specific feedback to therapists on the efficacy of their counseling. CORE-MI combines speech and language processing with automated coding and interactive visualization to help therapists and trainers identify strengths and areas for improvement, to help them provide effective counseling to their clients.

Author Keywords
Motivational interviewing, psychotherapy, substance abuse, patient-provider communication, speech recognition, language processing, automated coding, interactive visualization

ACM Classification Keywords
H.5.2 User Interfaces; J.3.c Medical information systems

INTRODUCTION
Millions of Americans utilize psychotherapy for substance abuse treatment [1]. Motivational Interviewing (MI) is a widely studied and effective method of treatment [2]. However, unlike pharmaceuticals where quality control can be

determined during manufacturing, psychotherapies like MI are primarily a conversation between the client (i.e., patient) and provider (i.e., counselor or therapist). Thus, quality cannot be determined until the interaction occurs.

In training contexts and clinical trials, the evaluation of psychotherapy providers relies on human judgment. Specifically, a human listens to a recording of a session and makes decisions regarding whether certain therapist behaviors have occurred. Feedback can be provided informally (e.g., a supervisor listens to the session and offers their perspective and input) or in the form of standardized, numeric ratings from a theory-derived rating system – a process that has remained fundamentally unchanged for decades [3]. Given the labor intensiveness of human ratings of psychotherapy, and the scale of demand in clinical practice, it is not surprising that most real-world psychotherapy is never evaluated. Thus, the quality of psychotherapy in clinical practice remains largely unknown, and providers rarely receive feedback on the quality of their care [4].

Kahneman and Klein [5] argue that specific, immediate feedback is critical to the development of expertise across a broad swath of domains and skills. Tracey et al. [6] note that specific, immediate feedback is precisely what is missing from the practice of psychotherapy, making the case that "expertise" in psychotherapy is challenging to attain. There is a strong need for an evaluation technology that makes feedback that is proximal to the clinical encounter feasible, which led to the development of the Counselor Observer Ratings Expert for MI (CORE-MI).

RELATED WORK
Visualizing Provider-Patient Communication
CORE-MI extends a growing body of HCI research on mediating and documenting communication between doctors and patients to improve medical care. It's

immediate predecessor is a system called Monologger, which provided report-card like feedback to oncologists [7]. Other related work includes providing patients with video recordings of medical visits [8], projecting images of anatomy on patients' bodies to aid understanding of medical problems in physical therapy [9], and visualizing nonverbal cues to enhance clinicians' perception of social signals [10]. CORE-MI is also informed by prior work on visualizing non-medical conversations to facilitate reflection [11] and track participation and topics in multi-person conversation [12].

Motivational Interviewing
CORE-MI provides feedback on Motivational Interviewing (MI), a type of substance abuse and health behavior change intervention that relies on empathic, collaborative, nonjudgmental communication. In MI, the therapist listens for and reinforces client "change talk" (i.e., talk about changing a problematic behavior), while not judging or confronting client "sustain talk" (i.e., talk about keeping a problematic behavior the same). There are a large number of clinical trials supporting the efficacy of MI, though research on training has shown that therapists' skill often decreases over time without ongoing performance-based feedback [13]. Below we offer an automated system that codes therapy sessions, and provides understandable, visual indication of provider fidelity.

CORE-MI
CORE-MI is an automated coding and visualization system that provides report-card like feedback on psychotherapists' adherence with MI. CORE-MI combines speech and language processing with automated coding and interactive visualization to help therapists and trainers identify strengths and areas for improvement, to help them provide effective counseling to their clients.

Speech and Language Processing
Although not the focus of the current paper, CORE-MI utilizes foundational speech signal processing methods to translate an audio-recording of an MI session into a numeric representation of semantic and vocal acoustic data, which are in turn used as features in machine learning predictive models. Key processing steps include: voice activity detection (i.e., at each tenth of a second, is someone speaking or not?), speaker segmentation (also called "diarization"; i.e., at each tenth of a second, which person is speaking?), role identification (i.e., *who* is speaking at each moment, counselor or client?), and automated speech recognition [14]. Additionally, paralinguistic information such as prosody, pitch, speech rate, and intensity are all estimated. A variety of machine learning approaches have been used to take these speech and language derived inputs and predict MI-relevant features of the session. CORE-MI makes use of the Barista open-source software platform [15] to provide an end-to-end system, with a JSON-based API to the CORE-MI report.

Automated Coding
CORE-MI provides feedback on standard MI quality measures described in the Motivational Interviewing Treatment Integrity Scale [16]. Two measures are considered "Global Ratings" and meant to capture the overall quality or "gestalt" of the session: MI Spirit and Empathy. MI Spirit assesses the overall competence of the counselor along dimensions of collaboration, evocation, and autonomy. In addition, Empathy measures the extent to which the therapist makes an effort to understand the client's perspective. Both overall measures are rated on a 1-5 likert scale.

CORE-MI also tracks three "Behavior Count" measures that characterize the quantity and quality of questions and reflections in the discourse. Questions and reflections are key discursive elements in MI. Behavior Count quality metrics include: 1) Reflection to Question Ratio, which captures the relative amount of reflective listening statements as compared to questions, 2) Percent Open Questions – open questions (as opposed to closed questions) are preferred as they allow a wide range of possible answers, and 3) Percent Complex Reflections – this is the percentage of all therapist reflective utterances that are "complex" rather than "simple". Complex reflections add meaning to client utterances or summarize what they client has said across previous statements, whereas simple reflections can be simple restatements.
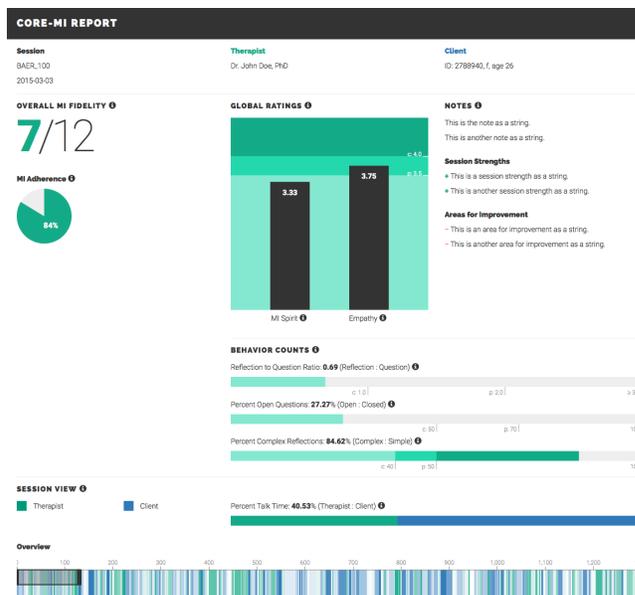

Figure 1: CORE-MI Report Visualization

Interactive Visualization
CORE-MI provides a report card-like, visual summary of counseling sessions **(figure 1)**. The report prominently features an "Overall MI Fidelity" score that aggregates all measures to give an impressionistic view of the participant's general level of adherence to MI principles. Bar charts indicate global and behavior count measures, highlighting the counselor's level of proficiency in MI techniques.

The visualization also offers a "Session View" that enables users to examine detailed session transcripts and talk turn level annotations for therapist and client behaviors (e.g., was the talk turn a question or

reflection, was the client talk-turn segment change or sustain talk). The session view illustrates turn taking, with therapist and client speech indicated in different colors. Color intensity is mapped to vocally-encoded arousal.

By selecting talk segments, a user can review the associated transcript text. In this way, the session view provides an overview of key session dynamics, directs the user's attention to areas of greatest interest, and enables her to review key exchanges.

The visualization is implemented in HTML and Javascript, and makes us of the D3 visualization library. It is accessible through standard web broswers on a variety of devices and platforms, but the design is optimized for tablet-like devices.

CURRENT STATUS
Technical feasibility of signal and language processing, and automated coding was previously reported [14]. At the time of this writing, visualization software has recently been developed, and a usability study is planned for Spring 2016. We anticipate reporting preliminary results at the Computing and Mental Health Workshop.

FUTURE WORK
Rapid, objective feedback is necessary for skill development and mastery, but thus far, this has not been possible with counseling or psychotherapy, given the reliance on human judgment for evaluation. To the authors' knowledge, the current work is the first effort to provide a technology for automating the evaluation of counseling and feedback to therapists. The CORE-MI summary report represents the critical front-end, user interface, which must provide an easy-to-understand summary of the evaluation metrics generated by the CORE-MI speech signal processing engine. Several future directions are planned for this work.

As noted earlier, a pilot usability study is beginning in February 2016. In this study participating therapists will have a brief, addiction-focused session with an actor portraying a client based on vignettes from real clients. Sessions will be recorded and feedback will be generated by the CORE-MI system from these recordings. In a follow-up session, therapists will provide quantitative and qualitative information about the acceptability of computer-generated feedback, ease of understanding the information presented, and satisfaction with the summary report. In addition, detailed questions will follow-up on each distinct data element presented in the CORE-MI report. This critical usability data will inform further revisions of the CORE-MI report.

A second future direction will be a clinical study of augmenting traditional MI training with feedback from the CORE-MI system. In this technology-scaffolded design, trainee therapists will not only receive in-person workshop-based training in MI, but following the training, CORE-MI summary reports on their sessions will help reinforce and remind trainees of the appropriate MI skills. This study will examine whether the CORE-MI augmented training enhances long-term skill retention of therapists as well as whether the technology in turn improves client outcomes.

Finally, the current implementation of the CORE-MI system focuses on individual sessions; however, typical clinical practice and training has a hierarchical structure. Individual therapists have a caseload of clients who come for therapy on repeated occasions. Thus, future work will extend the existing CORE-MI user-interface to a web-based dashboard design in which multiple sessions from a single patient may be summarized and compared, and similarly multiple clients from the same therapist may be summarized and compared.

CONCLUSIONS
Psychotherapy, and MI in particular is a primary treatment modality for substance use disorders. The CORE-MI offers the first scalable technology for evaluating the quality of client-therapist interactions. Moreover, even in well funded clinical trials where regular feedback on treatment quality is available, there may be little attention paid to the presentation of feedback (i.e., therapists may simply be presented with a list of scores). Here, the CORE-MI offers an additional contribution wherein feedback not only occurs rapidly, but is presented in a way that may facilitate therapist understanding and engagement with the raw material of the session. We are hopeful technological augmentations of psychotherapies similar to CORE-MI MI may lead to improvements in the quality of treatment available in the community and contribute to a reduction in the disease burden of substance abuse disorders.

REFERENCES
[1] SAMHSA, "The Treatment Episode Data Set (TEDS) 2000–2010. State Admissions to Substance Abuse Treatment Services.," Substance Abuse and Mental Health Services Administration, HHS Publication No. SMA-12-4729, 2012.
[2] W. Miller and S. Rollnick, *Motivational Interviewing*, 3rd ed. New York: Guilford, 2012.
[3] H. Kirschenbaum, "Carl Rogers's Life and Work: An Assessment on the 100th Anniversary of His Birth," *J. Couns. Dev.*, vol. 82, no. 1, pp. 116–124, Jan. 2004.
[4] S. Baldwin and Z. Imel, "Therapist Effects," in *Bergin and Garfield's Handbook of Psychotehrapy and Behavior Change*, 5th ed., M. Lambert, Ed. New Jersey: Wiley, 2013, pp. 258–297.
[5] D. Kahneman and G. Klein, "Conditions for intuitive expertise: a failure to disagree," *Am. Psychol.*, vol. 64, no. 6, pp. 515–526, Sep. 2009.
[6] T. J. G. Tracey, B. E. Wampold, J. W. Lichtenberg, and R. K. Goodyear, "Expertise in psychotherapy: an elusive goal?," *Am. Psychol.*, vol. 69, no. 3, pp. 218–229, Apr. 2014.

[7] J. Cook and T. Hirsch, "Monologger: visualizing engagement in doctor-patient conversation," 2014, pp. 37–40.

[8] P.-A. Hillgren and P. Linde, "Collaborative articulation in healthcare settings: towards increased visibility, negotiation and mutual understanding," 2006, pp. 301–310.

[9] T. Ni, A. K. Karlson, and D. Wigdor, "AnatOnMe: facilitating doctor-patient communication using a projection-based handheld device," 2011, p. 3333.

[10] R. A. Patel, A. Hartzler, M. P. Czerwinski, W. Pratt, A. L. Back, and A. Roseway, "Leveraging visual feedback from social signal processing to enhance clinicians' nonverbal skills," 2013, p. 421.

[11] P. Mathur and K. Karahalios, "Visualizing remote voice conversations," 2009, p. 4675.

[12] T. Bergstrom and K. Karahalios, "Social Mirrors as Social Signals: Transforming Audio into Graphics," *IEEE Comput. Graph. Appl.*, vol. 29, no. 5, pp. 22–32, Sep. 2009.

[13] C. S. Schwalbe, H. Y. Oh, and A. Zweben, "Sustaining motivational interviewing: a meta-analysis of training studies," *Addict. Abingdon Engl.*, vol. 109, no. 8, pp. 1287–1294, Aug. 2014.

[14] B. Xiao, Z. E. Imel, P. G. Georgiou, D. C. Atkins, and S. S. Narayanan, "'Rate My Therapist': Automated Detection of Empathy in Drug and Alcohol Counseling via Speech and Language Processing," *PLOS ONE*, vol. 10, no. 12, p. e0143055, Dec. 2015.

[15] D. Can, J. Gibson, C. Vaz, P. G. Georgiou, and S. S. Narayanan, "Barista: A framework for concurrent speech processing by USC-SAIL," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014): Florence, Italy, 4 - 9 May 2014*, Piscataway, NJ, 2014, pp. 3306–3310.

[16] T. B. Moyers, T. Martin, J. K. Manuel, W. R. Miller, and D. Ernst, "Revised Global Scales: Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1)." University of New Mexico Center on Alcoholism, Substance Abuse and Addictions (CASAA), 22-Jan-2010.