

PSYCHOTHERAPY RESEARCH¹

What About the Words? Natural Language Processing in Psychotherapy

Brian T. Pace, MS

Michael Tanana, M.Stat

Department of Educational Psychology
University of Utah, Salt Lake City



Bo Xiao, PhD

Department of Electrical Engineering
University of Southern California, Los Angeles



Aaron Dembe, BA

Christina S. Soma, BA

Department of Educational Psychology
University of Utah, Salt Lake City



Mark Steyvers, PhD

Department of Cognitive Sciences
University of California-Irvine

Shrikanth Narayanan, PhD

Department of Electrical Engineering
University of Southern California, Los Angeles

David C. Atkins, PhD

Department of Psychiatry and Behavioral Sciences
University of Washington, Seattle

Zac E. Imel, PhD

Department of Educational Psychology
University of Utah, Salt Lake City

By ‘augmenting human intellect’ we mean increasing the capability...to approach a complex problem situation...a way of life in an integrated domain where hunches, cut-and-dry, intangibles, and the human ‘feel for a situation’ usefully co-exist with powerful concepts, streamlined terminology and notation, sophisticated methods, and high-powered electronic aids. (Engelbart, 1962/2001, p.1)

Psychotherapy is certainly a complex situation, represented by the manifold of words, sounds, facial expressions, and body movements a therapist and client may present during a session. To help

clients, therapists rely on their own skill and training to manage a series of psychological processes that occur during a session—often isolated in their counsel-

continued on page 15

ing rooms with little external support (Miller, Sorensen, Selzer, & Brigham, 2006). Psychotherapists find themselves confronting concerns about variability in quality of care (Institute of Medicine, 2015), whether therapists improve with experience (Goldberg et al., 2016; Tracey, Wampold, Lichtenberg, & Goodyear, 2014), and the decline of psychotherapy as a percentage of mental health care (Olfson et al., 2002). Indeed, a recent article in *Science Magazine* questioned whether humans make good therapists at all, raising the specter that emerging natural language processing (NLP) technologies may one day improve upon and replace human therapists with computers (Bohannon, 2015)—a worrisome thought, but unlikely to occur anytime soon (Barrett & Gershkovich, 2014).

The most successful technological innovations in computer science are not those in which humans are replaced entirely, but instead involve the augmentation of human ability (Isaacson, 2014; Narayanan & Georgiou, 2013). However, since the 1940s and Carl Rogers' first recording of a psychotherapy session, little has changed in the evaluation of psychotherapy. Patients fill out self-report surveys and, at times, human coders are trained to rate therapy sessions, assessing adherence or competence in the use of a particular treatment (Imel, Steyvers, & Atkins, 2014). The practice of coding a psychotherapy session involves significant resources, including the cost of training and maintaining reliability amongst a team of coders assessing for theoretically driven criteria. Thus, outside of well-funded research settings, direct observation of psychotherapy is rare and often unfeasible.

But 70 years after Rogers' first recordings, and 50 years since Engelbart's seminal paper quoted above, we are on the cusp of major advances in psychother-

apy research and delivery. New methods are being developed that allow researchers to model the linguistic and semantic raw data of psychotherapy, which may improve both the specificity and scale of research on how treatments work. This work may generate new approaches to providing process feedback to therapists. Indeed, rather than replace humans, technology may soon make it possible to provide therapists and patients rapid, objective feedback on treatment, and conduct large-scale mechanism analyses of data from thousands—and eventually millions—of therapy sessions.

What Is Natural Language Processing?

NLP is a subfield of computer science and machine learning where the goal is to “learn, understand, and produce human language content” (Hirschberg & Manning, 2015, p. 261). In essence, NLP methods take large collections of unstructured text as inputs and generate more useful information. For example, NLP might use transcripts from psychotherapy sessions to help answer questions like: “What were they talking about?,” “Is this person depressed?,” or “Is this therapist empathic?” With NLP, large text corpora, that were previously unusable without human interpretation and judgment, become accessible rich data sources. Some of the recent successes of NLP include sentiment analysis: NLP models can now identify whether complicated statements are positive or negative, almost as well as humans (Socher et al., 2013). These models can detect whether sentences are paraphrases of each other, and can translate between different languages (Socher, Huang, Pennington, Ng, & Manning, 2011; Bahdanau, Cho, & Bengio, 2015). NLP models can summarize large documents and classify the topic of an article (Steyvers & Griffiths, 2007), or identify

continued on page 16

authorship (Pearl & Steyvers, 2012; Zhao & Zobel, 2005). More recent NLP models attempt to produce language and dialogue (Vinyals & Le, 2015). In mental health, researchers have made progress toward using individuals' Twitter feeds to identify those who are depressed or later go on to be diagnosed with schizophrenia (Mitchell, Hollingshead & Coppersmith, 2015; Mowery, Way, Bryan & Conway, 2015).

Psychotherapy and NLP

Psychotherapy typically involves a conversation, and conversations contain an abundance of words. A typical 50-minute session may include about 12,000 to 15,000 words (Lord et al., 2015). In a small clinical trial (say 10 sessions x 20 patients) there could be 2.7 million spoken words. Thus, similar to other large text corpora, NLP provides a platform to analyze this text, using methods that do not rely solely on labor intensive human coding. Below, we describe several examples from our own work, and several from others, in which NLP methods are utilized to evaluate psychotherapy, including: a) testing theoretical models of emotional/relational processes; b) exploring content and symptoms discussed; c) categorizing treatment and utterance level coding of provider fidelity to treatments; and d) fully automatic rating and feedback to providers directly from session audio.

Relational/emotional processes. A primary focus of the early use of NLP methods in psychotherapy has been to evaluate complex relational/emotional processes (e.g., empathy) using the words from treatment sessions. Much of this work has involved the use of computerized dictionaries that place specific words in psychologically meaningful categories (e.g., emotion words, reflecting or experiencing; e.g., Mergenthaler, 2008). For example, Anderson and colleagues (1999) found that when the patient used

more emotion words, therapists obtained better outcomes when minimizing responses with cognitively geared verbs (e.g., think, believe, know). Using a similar technology, Mergenthaler and colleagues have assessed temporal trends in emotional tone (i.e., words that indicate emotional relevant language) and abstraction (i.e., "conceptual language") across the course of a therapy session (Buchheim & Mergenthaler, 2000; Mergenthaler, 2008, p. 113; Mergenthaler, 1996).

More recently, we tested a hypothesis about the development of empathic synchrony (Preston & de Waal, 2002), using therapist and client linguistic style synchrony (LSS) in neighboring talk turns (here LSS implies a matching of specific word tokens in therapist-client phrases). LSS was significantly higher in sessions rated by humans as high versus low empathy (Lord, Sheng, Imel, Baer, & Atkins, 2015). However, as the categories in these NLP programs are defined by humans, a primary limitation is that the computer cannot "learn" underlying structure from the data. For example, these models can struggle with polysemy (i.e., words can have multiple meanings depending on context), and thus the word "like" may erroneously fall into a positive emotion category, when it is being used as filler in the sentence (e.g., "Like, how are you doing?").

Exploring content and classifying interventions. As human methods for examining the content of psychotherapy are so labor intensive, it is relatively rare to conduct large scale explorations of what therapists and clients talk about. In a more recent study, we used an NLP method called a topic model (Steyvers & Griffiths, 2007) to classify the content of 1,533 psychotherapy or pharmacotherapy session transcripts—1.2 million

continued on page 17

words and 223,000 patient-therapist talk turns (Imel, Steyvers, & Atkins, 2014; see Atkins et al., 2012 for a tutorial on topic models and a similar example in couples therapy). The model identified specific semantically relevant topics (e.g., the topic “depression” included word tokens, such as self, fine, sad, hopeless, appetite, helpless, and esteem). We were able to use these session level topic labels to identify specific talk turns where that topic occurred (Gaut, Steyvers, Imel, Atkins, & Smyth, 2015). In addition, we used a “labeled” topic model (i.e., a semi-supervised topic model; Rubin, Chambers, Smyth, & Steyvers, 2012) to learn the language associated with a particular therapeutic label. For example, the following was predicted to be an utterance related to Cognitive Behavioral Therapy (CBT): “To succeed. That is kind of your main or irrational belief. ‘I should not have to work as hard as other people to succeed.’” The model also automatically assigned sessions to one of four types of treatment (i.e., CBT, Humanistic/Experiential, Psychodynamic, and Medication Management) with high accuracy (Imel et al., 2014).

Utterance Level Coding of Transcripts

Perhaps the most time demanding task in coding psychotherapy sessions is the painstaking process of assigning unique labels to each and every utterance that occurs in a session—essentially reducing the mass of words in a transcript to a reduced set of psychologically meaningful labels. For example, a recent meta-analysis (Magill et al., 2014) of the literature focusing on treatment mechanisms in Motivational Interviewing (MI) included 0.3% of MI sessions included in clinical trials (Lundahl, Kunz, Brownell, Tollefson, & Burke, 2010). A major focus of our team’s work has been to train various NLP models to annotate transcripts based on the words in a therapist or client utterance. In an initial paper, we used a

version of the topic model noted above to predict utterance level Motivational Interviewing Skills Codes (MISC; Miller, Moyers, Ernst, & Amrhein, 2008; e.g., reflections, questions, etc.) in 155 MI sessions (Atkins, Steyvers, Imel, & Smyth, 2014). In a larger follow-up study ($n = 341$ sessions; 78,977 patient-provider talk turns), we used other text-based models that considered the syntax and semantics of a statement [i.e., (a) discrete sentence features and (b) recursive neural networks; Tanana, Hallgren, Imel, Atkins & Srikumar, in press]. We identified closed and open questions, simple and complex reflections, affirmations, and giving information behavioral codes as well or better than human coders (see <http://sri.utah.edu/psychtest/> for testing the models on your own text and correct errors; Tanana et al., in press).

From sound to codes. A limitation of the above NLP work is that the models require transcripts, limiting the potential of NLP models to scale up to larger tasks. Automatic speech recognition (ASR; the technology used in applications like Siri or Cortana) is a method where a speech recognition system transcribes spoken content (see Hinton et al., 2012, for a recent implementation). In a fully automated system that combined ASR with an NLP prediction model for therapist empathy, we found computer based empathy ratings starting from audio recordings were strongly correlated ($r = 0.62$) with observed, human-rated sessions (Xiao, Imel, Georgiou, Atkins & Narayanan, 2015). In a follow-up using the same algorithm, we role-played 20 MI sessions in the lab. These sessions were split between “good” and “bad” MI therapy conditions, where the counselors in the former condition intentionally emphasized MI adherent counseling skills and the latter emphasized MI non-adherent skills (see Figure

continued on page 18

1). Even in this new data, human and machine scores were strongly correlated ($r = 0.61$), and machine-generated scores were clearly differentiated across the two therapy role play conditions, $t(18) = 6.2, p < .001, d = 2.9$ (see Xiao, Huang, Imel, Atkins, Georgiou, & Narayanan, under review, for a more detailed technical description of the system).

reflections), but will also include a detailed session view in which the counselor can examine ASR-generated text from the entire session and associated behavioral codes, changes in emotional arousal, and talk time. Of course, this work need not be focused on ML, and we hope it will soon be adapted to a variety of interventions and contexts.

Figure 1: Human Versus Machine Empathy Ratings

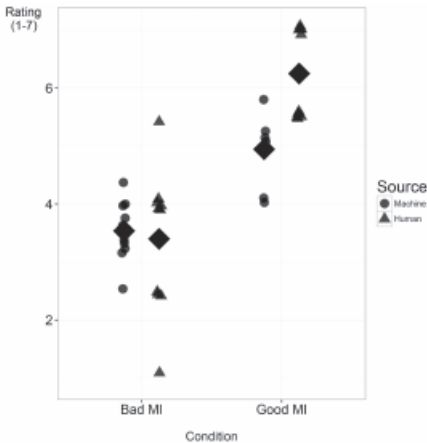


Figure 1. A comparison of human versus machine empathy ratings of 20 role-played “good” and “bad” MI therapy sessions. The machine ratings are represented by circles and the human ratings by triangles. Each circle or triangle is a single session rating. The diamonds represent the mean rating for each category.

The Future

In this brief review, we have focused on one technical solution to a problem that limits progress in psychotherapy science and practice—namely a need for scalable tools that can evaluate what occurs during the treatment hour. At present we are beginning a National Institute of Health (NIH) funded usability study that uses the “sound to codes” infrastructure highlighted above, to provide rapid feedback to counselors on their use of MI. The provider feedback is a web-based tool with ratings on standard MI fidelity measures (e.g., empathy and

In the last decade, we have seen a revolution in the use of practice-based evidence in psychotherapy. As soon as it became possible to measure client outcomes in large-scale databases, it also became possible to study variability in dosage, response, and therapist-to-therapist variability (e.g., Stiles, Barkham, Wheeler, 2015; Wampold & Brown, 2005). However, that data leaves us wondering why some therapists are “good” and why others are “bad.” ASR and NLP tools provide the framework to understand critically important psychotherapy processes on a previously impossible scale. A future possibility is that NLP modeling could focus on outcomes (rather than treatments), and attempt to identify the linguistic features of treatments characterized by “good” and “bad” treatment outcomes. These results could inform practice and generate feedback to therapists. As such, this technology may open the door to explore the use of highly detailed, practice-based evidence to inform evidence-based practice. It is our hope that this sort of detailed feedback could one day be available to individual therapists to promote reflective practice and facilitate ongoing development of expertise, ultimately reducing the suffering of the patients we aim to help.

References for this article can be found in the online version of the Bulletin published on the Society for the Advancement of Psychotherapy website.



-
- highly experienced therapists. Paper presented at the 45th Society for Psychotherapy Research International Annual Meeting, Copenhagen, Denmark.
- Råbu, M., Binder, P., & Haavind, H. (2013). Negotiating ending: A qualitative study of the process of ending psychotherapy. *European Journal of Psychotherapy & Counselling, 15*(3), 274-295. doi: 10.1080/13642537.2013.810962
- Råbu, M., Haavind, H., & Binder, P. (2013). We have travelled a long distance and sorted out the mess in the drawers: Metaphors for moving towards the end in psychotherapy. *Counselling and Psychotherapy Research, 13*(1), 71-80. doi: 10.1080/14733145.2012.711339
- Råbu, M., Halvorsen, M. S., & Haavind, H. (2011). Early relationship struggles: A case study of alliance formation and reparation. *Counselling and Psychotherapy Research, 1*(1), 23-33. doi: 10.1080/14733145.2011.546073
- Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational Behavior and Human Decision Processes, 53*(2), 252-266. doi: 10.1016/0749-5978(92)90064-E
- Shanteau, J., & Weiss, D. J. (2014). Individual expertise versus domain expertise. *American Psychologist, 69*(7), 711-712. doi: 10.1037/a0037874
- Tracey, T. J. G., Wampold, B. E., Goodyear, R. K., & Lichtenberg, J. W. (2015). Improving expertise in psychotherapy. *Psychotherapy Bulletin, 50*(1), 7-13.
- Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal. *American Psychologist, 69*, 218-229. doi: 10.1037/a0035099
- von der Lippe, A., Oddli, H. W., & Halvorsen, M. S. (2014, June). *What do therapists do who succeed with difficult to treat clients?* Paper presented at the 45th Society for Psychotherapy Research International Annual Meeting, Copenhagen, Denmark.
- What About the Words? Natural Language Processing in Psychotherapy**
- Anderson, A., Bein, E., Pinnell, B. J., & Strupp, H. H. (1999). Linguistic analysis of affective speech in psychotherapy: A case grammar approach. *Psychotherapy Research, 9*(1), 88-99. doi: 10.1093/ptr/9.1.88
- Atkins, D. C., Rubin, T. N., Steyvers, M., Doeden, M. A., Baucom, B. R., & Christensen, A. (2012). Topic models: A novel method for modeling couple and family text data. *Journal of Family Psychology, 26*(5), 816-827. doi:10.1037/a0029607
- Atkins, D. C., Steyvers, M., Imel, Z. E., & Smyth, P. (2014). Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science, 9*(1), 49. doi: 10.1186/1748-5908-9-49
- Bahdanau, D., Cho, K., & Bengio, Y. (May, 2015). Neural machine translation by jointly learning to align and translate. Proceedings from the: International Conference on Learning Representations (ICLR), San Diego, CA.
- Barrett, M. S., & Gershkovich, M. (2014). Computers and psychotherapy: Are we out of a job? *Psychotherapy, 51*(2), 220-223. doi: 10.1037/a0032408
- Bohannon J. (2015). The synthetic therapist: Some people prefer to bare their souls to computers rather than to fellow humans. *Science, 349*(6245), 250-251. doi: 10.1126/science.349.6245.250.
- Buchheim, A., & Mergenthaler, E. (2000). The relationship among attachment representation, emotion-abstraction patterns, and narrative
-

-
- style: A computer-based text analysis of the adult attachment interview. *Psychotherapy Research*, 10(4), 390–407.
- Engelbart, D. C. (2001). Augmenting human intellect: A conceptual framework (1962). In R. Packer & K. Jordan (Eds.), *Multimedia: From Wagner to Virtual Reality* (pp. 64–90). New York, NY: WW Norton & Company.
- Gaut, G., Steyvers, M., Imel, Z., Atkins, D., & Smyth, P. (2015). Content coding of psychotherapy transcripts using labeled topic models. *IEEE Journal of Biomedical and Health Informatics*, 1–12. doi: 10.1109/JBHI.2015.2503985
- Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E. (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology*, 63(1), 1–11. doi: 10.1037/cou0000131
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., ... Kingbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6), 82–97.
- Hirschberg J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261–266. doi: 10.1126/science.aaa8685
- Imel, Z. E., Steyvers, M., & Atkins, D. C. (2014). Computational psychotherapy research: Scaling up the evaluation of patient–provider interactions. *Psychotherapy*, 52(1), 19–30. doi: 10.1037/a0036841
- Institute of Medicine (IOM). (2015). *Psychosocial interventions for mental and substance use disorders: A framework for establishing evidence-based standards*. Washington, DC: The National Academies Press.
- Isaacson, W. (2014). *The innovators: How a group of hackers, geniuses, and geeks created the digital revolution*. New York, NY: Simon & Schuster.
- Lord, S. P., Can, D., Yi, M., Marin, R., Dunn, C. W., Imel, Z. E., ... Atkins, D. C. (2015). Advancing methods for reliably assessing motivational interviewing fidelity using the motivational interviewing skills code. *Journal of Substance Abuse Treatment*, 49, 50–57. <http://doi.org/10.1016/j.jsat.2014.08.005>
- Lord, S. P., Sheng, E., Imel, Z. E., Baer, J., & Atkins, D. C. (2015). More than reflections: Empathy in motivational interviewing includes language style synchrony between therapist and client. *Behavior Therapy*, 46(3), 296–303. doi: 10.1016/j.beth.2014.11.002
- Lundahl, B. W., Kunz, C., Brownell, C., Tollefson, D., & Burke, B. L. (2010). A meta-analysis of motivational interviewing: Twenty-five years of empirical studies. *Research on Social Work Practice*, 20(2), 137–160. doi: 10.1177/1049731509347850
- Magill, M., Gaume, J., Apodaca, T. R., Walthers, J., Mastroleo, N. R., Borsari, B., & Longabaugh, R. (2014). The technical hypothesis of motivational interviewing: A meta-analysis of MI's key causal model. *Journal of Consulting and Clinical Psychology*, 82(6), 973–983. doi: 10.1037/a0036833
- Mergenthaler, E. (1996). Emotion–abstraction patterns in verbatim protocols: A new way of describing psychotherapeutic processes. *Journal of Consulting and Clinical Psychology*, 64(6), 1306–1315. doi: 10.1037/0022-006X.64.6.1306
- Mergenthaler, E. (2008). Resonating minds: A school-independent theoretical conception and its empirical application to psychotherapeutic processes. *Psychotherapy Research*, 18(2), 109–125. doi: 10.1080/10503300701883741
- Miller, W. R., Moyers, T. B., Ernst, D., & Amrhein, P. (2008). *Manual for the Motivational Interviewing Skill Code*

-
- Version 2.1 (MISC). Unpublished manual. University of New Mexico, Albuquerque, NM.
- Miller, W. R., Sorensen, J. L., Selzer, J. A., & Brigham, G. S. (2006). Disseminating evidence-based practices in substance abuse treatment: A review with suggestions. *Journal of Substance Abuse Treatment, 31*(1), 25–39. doi: 10.1016/j.josat.2006.03.005
- Mitchell, M., Hollingshead, K., & Coppersmith, G. (2015). Quantifying the language of schizophrenia in social media. *Proceedings from the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, NAACL-CLPsych, 11-20*.
- Mowery, D. L., Way, W., Bryan, C., & Conway, M. (2015). Towards developing an annotation scheme for depressive disorder symptoms: A preliminary study using twitter data. *Proceedings from the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, NAACL-CLPsych, 89–98*.
- Narayanan, S., & Georgiou, P. G. (2013). Behavioral signal processing: Deriving human behavioral informatics from speech and language. *Proceedings from the Institute of Electrical and Electronics Engineers (IEEE), 101*(5), 1203–1233. doi: 10.1109/JPROC.2012.2236291
- Olson, M., Marcus, S. C., Druss, B., Elinson, L., Tanielian, T., & Pincus, H. A. (2002). National trends in the outpatient treatment of depression. *Journal of American Medical Association, 287*(2), 203–209.
- Pearl, L., & Steyvers, M. (2012). Detecting authorship deception: A supervised machine learning approach using author writeprints. *Literary and Linguistic Computing, 27*(2), 183–196. doi: 10.1093/lc/fqs003
- Preston, S. D., & de Waal, F. B. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences, 25*(1), 1–20. doi: 10.1017/S0140525X02000018
- Rubin, T. N., Chambers, A., Smyth, P., & Steyvers, M. (2012). Statistical topic models for multi-label document classification. *Machine Learning, 88*(1-2), 157–208. doi: 10.1007/s10994-011-5272-5
- Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. Proceedings of the Conference on Empirical Methods, Seattle, WA.
- Socher, R., Huang, E., Pennington, J., Ng, A. Y., & Manning, C. (2011). Dynamic pooling and unfolding recursive autoencoders for paraphrase detection. *Advances in Neural Information Processing Systems, 1–9*.
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis*, (424–448). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Stiles, W. B., Barkham, M., & Wheeler, S. (2015). Duration of psychological therapy: Relation to recovery and improvement rates in UK routine practice. *The British Journal of Psychiatry, 207*(2), 115–122. doi: 10.1192/bjp.bp.114.145565
- Tanana, M., Hallgren, K., Imel, Z. E., Atkins, D. C., & Srikumar, V. (in press). A comparison of natural language processing methods for automated coding of motivational interviewing. *Journal of Substance Abuse Treatment*.
- Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist, 69*(3), 218–229. doi: 10.1037/a0035099
- Vinyals, O., & Le, Q. V. (2015). A neural conversational model. *Proceedings of*
-

the 31st International Conference on Machine Learning, *Journal of Machine Learning Research*, 37.

Wampold, B. E., & Brown, G. S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology*, 73(5), 914–923. doi: 10.1037/0022-006X.73.5.914

Xiao, B., Imel, Z. E., Georgioui, P. G., Atkins, D. C., & Narayanan, S. S. (2015). "Rate my therapist": Automated detection of empathy in drug and alcohol counseling via speech and language processing. *Public Library of Science One*, 10(12), 1-15. doi: 10.1371/journal.pone.0140555

Xiao, B., Huang, C. W., Imel, Z. E., Atkins, D. C., Georgioui, P., & Narayanan, S. (2015). A technology prototype system for rating therapist empathy from audio recordings in addiction counseling. Manuscript submitted for publication.

Zhao, Y., & Zobel, J. (2005). Effective and scalable authorship attribution using function words. *Information Retrieval Technology*, 3689, 174–189. doi: 10.1007/11562382_14

The Role of Psychotherapists in the Training of Master's Level Counselors

American Psychological Association (APA). (2016). Commission on accreditation overview. Retrieved from: <http://www.apa.org/ed/accreditation/about/coa/>

Anderson, T., Crowley, M. E., Himawan, L., Holmberg, J. K., & Uhlin, B. D. (2015). Therapist facilitative interpersonal skills and training status: A randomized clinical trial on alliance and outcome. *Psychotherapy Research*, 7, 1-19.

Council for Accreditation of Counseling and Related Educational Programs (2015). CACREP 2016 standards. Retrieved from <http://www.cacrep.org/wp-con->

<tent/uploads/2016/02/2016-Standards-with-Glossary-rev-2.2016.pdf>.

Falender, C. A., Cornish, J. A. E., Goodyear, R., Hatcher, R. Kaslow, N. J., Leventhal, G. ... Sigmon, S. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology*, 60(7), 771-785. doi: 10.1002/jclp.20013

Goldberg, S. B., Rousmaniere, T., Miller, S. D., Whipple, J., Nielsen, S. L., Hoyt, W. T., & Wampold, B. E., (2016). Do psychotherapists improve with time and experience? A longitudinal analysis of outcomes in a clinical setting. *Journal of Counseling Psychology*, 63(1), 1–11. doi: 10.1037/cou0000131

Okiishi, J. C., Lambert, M. J., Eggett, D., Nielsen, L., Dayton, D. D., & Vermeersch, D. A. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists and on their clients' psychotherapy outcome. *Journal of Clinical Psychology*, 62(9), 1157-1172. doi: 10.1002/jclp.20272

Practice Oriented Research: Clinical Implications and Benefits

Adelman, R. W., Castonguay, L. G., Kraus, D. R., & Zack, S. E. (2015). Conducting research and collaborating with researchers: The experience of clinicians in a residential treatment center. *Psychotherapy Research*, 25(1), 108-120. doi: 10.1080/10503307.2014.935520

Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, 25(1), 6–19. doi:10.1080/10503307.2013.81769

Castonguay, L. G. (2011). Psychotherapy, psychopathology, research and practice: Pathways of connections and integration. *Psychotherapy Research*, 21(2), 125-140. doi: