HHS Public Access
Author manuscript
Interspeech. Author manuscript; available in PMC 2023 January 25.

Published in final edited form as:
Interspeech. 2019 September ; 2019: 1901–1905. doi:10.21437/interspeech.2019-2829.

Identifying Therapist and Client Personae for Therapeutic Alliance Estimation

Victor R. Martinez[1], **Nikolaos Flemotomos**[1], **Victor Ardulov**[1], **Krishna Somandepalli**[1], **Simon B. Goldberg**[2], **Zac E. Imel**[3], **David C. Atkins**[4], **Shrikanth Narayanan**[1]

[1]Signal Analysis and Interpretation Lab, University of Southern California, Los Angeles, CA, USA

[2]Department of Couseling Psychology, University of Wisconsin-Madison, Madison, WI, USA

[3]Department of Educational Psychology, University of Utah, Salt Lake City, UT, USA

[4]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, USA

Abstract

Psychotherapy, from a narrative perspective, is the process in which a client relates an on-going life-story to a therapist. In each session, a client will recount events from their life, some of which stand out as more significant than others. These significant stories can ultimately shape one's identity. In this work we study these narratives in the context of therapeutic alliance—a self-reported measure on the perception of a shared bond between client and therapist. We propose that alliance can be predicted from the interactions between certain types of clients with types of therapists. To validate this method, we obtained 1235 transcribed sessions with client-reported alliance to train an unsupervised approach to discover groups of therapists and clients based on common types of narrative characters, or *personae*. We measure the strength of the relation between personae and alliance in two experiments. Our results show that (1) alliance can be explained by the interactions between the discovered character types, and (2) models trained on therapist and client personae achieve significant performance gains compared to competitive supervised baselines. Finally, exploratory analysis reveals important character traits that lead to an improved perception of alliance.

Index Terms:

narratology; therapeutic alliance; persona

1. Introduction

In psychotherapy, clients seek to overcome a particular difficulty or problem with the help of a professional. It is uncommon for a client to know exactly what the root cause of their adversity really is. Instead, clients usually begin a therapy session by recounting a life-story related to the problems they want to work through, the significance and feelings they have associated with these events, and how they relate to their relationships with

victorrm@usc.edu, shri@sipi.usc.edu.

others. Some of these stories will stand out as more significant than others, usually the ones stemming from negative events. These significant stories can ultimately shape one's identity [1]. Professional therapists unravel a client's complex narratives as to understand what is of interest, and to raise awareness on particular traits and characteristics that may have caused the client's current afflictions. Moreover, therapists might also benefit in telling a story of their own. For example, they might build rapport with their client by recounting of a similar life-experience, or deliver a therapeutic narrative focused on the client—that is, when the therapist's story is really a retelling of the client's story. Hence, understanding the narrative elements of these stories, and the similarities between their stories, might help provide insights into how therapy works.

In this work we introduce a novel approach towards analyzing therapeutic outcomes from the stories told during psychotherapy sessions. Several factors contribute to positive therapeutic outcomes, some of which are strongly related to the combination of individuals and their relationship [2, 3, 4, 5]. One specific relationship factor, known as *therapeutic alliance* [6], corresponds to the collaborative aspects of the therapistclient relationship including the perception of a shared bond and the agreement on the focus of the therapy treatment. This factor is a major contributing element in psychotherapy success [2, 6, 7]. Unlike other counting measurements (e.g., ratio of open and closed questions [8]), it is unclear how alliance might be captured by what is discussed in each session [5, 2]. Instead, automatically assessment of therapeutic alliance might require higher-order cognitive and affective models for the individuals. To address this limitation, we propose to model alliance as a function on the interaction between certain types of therapists with certain types of clients. These types are automatically discovered from therapy transcripts by identifying the attributes shared among the client's and therapist's characters in the stories told throughout the sessions. Character's attributes are obtained through a Personae model [9]. Personae, also known as character archetypes, are classes of characters grouped by similar traits, behaviors and motivations [10]. For example, common personae in story-telling include *The Hero*, *The Villain* and *The Wise Old-Man*. While these models have found success in narrative understanding [9, 11], to the best of our knowledge no one has investigated their application on real-life narratives. Our approach is as follows: In each session, client and therapists will tell stories, and as with other stories, these narratives contain characters, setting, plot, conflicts, and resolution. The characters of those stories are imperfect portraits of persons in the client's life, which include both the client themselves (as the protagonist) and the therapist. For the client and therapist characters, we automatically identify their personae by using an unsupervised model for narrative understanding [9]. This model extends on the ideas of unsupervised topic modelling, in which documents (i.e. therapy sessions) are represented by mixtures of their characters' archetypes, and each archetype corresponds to a mixture of topics. To evaluate our approach, we first train a Persona model with automatically transcribed sessions. Then we analyze the relation between the interactions of characters' personae and alliance to show that the discovered personae is useful for the task of alliance estimation. Finally, we identify the most important character traits which promote the perception of alliance. To the best of our knowledge, this is the first work to connect the narrative processes underlying psychotherapy to alliance which measures the impact (success) of an intervention.

2. Previous Work

Narratives describe complex chains of events, with relationships between characters and objects that reflect the full complexity of life [12]. These are often communicated via sophisticated forms of discourse, relying on deep knowledge of the world, society, and culture [12, 13]. As a way to explore a person's internal state, narrative processes have been studied for mental health [14], medical interviews [15, 16], and children's language therapy [17]. In psychotherapy, narratives are used to allow clients to distance themselves from particular issues, in order to gain a new and objective perspective into the problem [1]. For example, in a form of counseling called narrative therapy [18], therapists and clients build upon storylines based on the client's dreams, values, goals and skills [1]. These storylines uncover the true nature of a client, separate from their problems. However, its effectiveness has only been supported by case studies and qualitative research, lacking clinical and empirical support found across other psychotherapy methods [19, 20]. In trying to understand the different aspects of a narrative, most of the computational work focuses on the sequence of events in which a story is defined [9]. This can be done either with a generative process [21] or with information extraction, through unsupervised topic discovery [22]. Some works have studied the characters in a narrative, their traits and behaviors—mostly for movies and novels [9, 23, 11].

To predict therapy effectiveness, many psychology studies have focused on understanding the client-therapist relation. Similarities between the therapist and client personalities have been associated with longer sessions, higher therapeutic alliance and overall therapy outcome [7, 24]. Project MATCH [25] studied if the outcome of treatments for alcoholism were improved by selecting type of treatment depending on a patient's characteristics. The authors' results suggest that that there is no benefit in matching patients to types of treatments. However, critics have pointed out that while treatment type was not significant, the way particular therapists interacted with alcoholics had a substantial impact on patient outcomes [26, 27].

Several methods for automatic assessment of treatment effectiveness have been proposed. Most of them rely on audio and linguistic cues (e.g., empathy and behavioral codes in therapy [28, 29]). Other methods explored the use of unsupervised topic modelling techniques as a higher-level measure of content, specifically their relation to mental health outcomes [30, 31]. In this work, we extend the study of narratives, from a character-centered perspective, to the context of psychotherapy. In particular, we explore the application of an unsupervised method to automatically discover classes of characters for both therapist and client, and investigate how the interactions between these classes predict therapy outcomes.

3. Data

For this study, we obtained 1235 recorded sessions of dyadic interactions between a therapist and a client. Sessions were collected between September 2017 and December 2018 at a US university counseling center. Explicit consent from both the therapist and clients was obtained before recording. The total number of clients and therapists is 386 and 40, respectively. These average duration of the sessions was 50.71 ± 10.32 minutes. Table 1 summarizes the distribution of the number of sessions per therapist and client. Before the start of each session except the first one, self-reported alliance is collected from the clients. Therapeutic alliance is scored along 4 dimensions using the short form of the Working Alliance Inventory[32, 33]. This form includes four items: "[*therapist*] and I are working towards mutually agreed upon goals", "I believe the way we are working on my problem is correct", "I feel that [*therapist*] appreciates me", and "[*therapist*] really understands me". Each item is scored by the client on a 7-point Likert scale and the average is used as the final alliance rating. A histogram of the alliance ratings in the available data is shown in Figure 1.

4. Method

As previously stated, we aim to model alliance as a function of the interaction between client's and therapist's types. As an indirect way of inferring a therapist's (client's) traits and motivations we present a method based on computational methods for persona discovery. These persona induce a natural clustering of characters, which we show can be then used to estimate their reported alliance more accurately than linguistic-based methods.

First, sessions were automatically transcribed using an speech processing pipeline. This pipeline is based on *state-of-the-art* models offered by Kaldi [34]. It consists of four steps: (1) Voice Activity Detection (VAD), (2) Diarization, (3) Automatic Speech Recognition (ASR), and (4) Role assignment. For VAD, a two-layer feed forward network with a softmax inference layer at the frame level was used. Diarization was based on the x-vector/PLDA paradigm [35]. For ASR, a time-delay neural network [36] and a tri-gram language model were trained on more than 4,000 hours of data from publicly available speech corpora augmented with noise and reverberation. We also adapted it using in-domain psychotherapy data. For role assignment, the two diarized speaker clusters were assigned to either therapist or client roles, following the method presented in [37]. Performance of this system was evaluated on additional psychotherapy sessions provided by the counseling center. Unweighted Average Recall for VAD was 82.7%, Diarization Speaker Error Rate was 6.4% and ASR Word Error Rate was 36.4%. The resulting ASR transcripts were processed for lemmatization, dependency parsing, and co-reference resolution using CoreNLP [38]. As a last step, we split each transcription by their speaker into two different documents.

Character identification.

Narratives may contain any number of characters, however we are only interested in those characters that represent the therapist and the client. To identify when a participant narrates actions corresponding to one of these two characters, we rely on the following assumption: the therapist (client) refers to their own character using first person pronouns only; conversely, they refer to the other participant's character with second person pronouns. For example, when narrating a story about the client's character, the therapist uses second person singular pronouns (i.e., "you") where the client refers to that same character using a 1st person singular pronouns (i.e., "I"). This process yields four different character combinations: (1) therapist character from therapist text, (2) therapist character from client text, (3) client character from therapist text, and (4) client character for client text. We explore the contribution of each one of these combinations as part of our experiments.

Personae model.

For each session, we assign a persona to each of the two characters. Each persona is selected from a distribution learned from the data using the unsupervised topic modelling technique first proposed by [9]. This model extends Latent Dirichllet Allocation [39] to narratives by assuming each story is generated by a mixture of characters' traits. These traits aim to capture the way in which a character is revealed through the narrative: by the actions they take toward others, the actions done to them, and the attributes used to describe them. Following this idea, a *persona* is represented as a triplet of multinomial distributions over the topics, capturing the action verbs, possessives, and modifiers. Each topic is represented as a weighted distribution over the complete vocabulary. Given a set of documents, Personae models learn P persona representations from a group of K topics where P and K are hyper-parameters. For this work, we applied the Personae model on the co-referenced transcriptions (see Figure 2). In each session, we extract persona representations for client and therapist transcriptions independently. This enforces the notion of different personae for different roles, and allows our evaluation to better weight on the interaction between therapists and client types. The participants are assigned to a single persona corresponding to the maximum posterior probability.

5. Evaluation

We use linear mixed effect models (LMEs) to measure the effect of characters' persona on therapeutic alliance. We fit an LME with alliance as response variable, and therapist and client personae as fixed effect variables. Additionally, we control for the therapist and client by including their unique anonymized identifiers as random effects. We experiment both with and without interaction terms. Quality of models is evaluated using Akaike Information Criterion (AIC), a measure based on information theory that rewards models on their goodness-of-fit while penalizing for their complexity. Models are compared against the null model (i.e., therapist and client identifiers only), and against models with varying number of topics ($K \in \{10, 20, 30, 50\}$) and number of personae ($P \in \{5, 10, 20, 30\}$). These comparisons are done using likelihood tests, correcting for multiple tests using Holm-Bonferroni method. For our experiments, we split our dataset into train (85%, $n = 1049$) and dev (15%, $n = 186$). Best values for K and P correspond to the model with the lowest value of AIC. However, this pair need not be unique, as it is possible for models with different K and P to perform just as well as the model with the best pair. To identify the minimum set of parameters that give rise to the best models, we follow the steps suggested by [40].

Regression experiments.

Additionally, we train machine learning models to capture the relation between alliance and assigned personae. As mentioned before, each persona consists of a triplet of multinomial distributions over action verbs, possessives and modifiers. We concatenate these distributions into a single vector to obtain a vectorial representation for the personas. We train support vector regressor (SVR) with a linear kernel on vector representations of the personae to predict alliance. We chose the linear kernel since it performed better in our preliminary experiments, and it allows us to identify the components that are most *important* for predicting alliance. Model performance is estimated using mean squared error (MSE) with cross-validation in a leave-one-therapist-out fashion. We compare the performance of our model to SVRs trained using uni- and bi-gram language models from either participant speech as well as trained on their joint text. These baselines were selected due to their success in related tasks [28].

6. Results

Model selection.

Our results show that regardless of the choice of K and P, models with only one of the therapist or client personae do not perform significantly better than the null model ($AIC(null)$ = 1091.75, $AIC(therapist)$ = 1088.75, $AIC(client)$ = 1095.90, χ^2 tests, all $p > 0.05$). In contrast, including both therapist and client personae significantly increases the explanatory power of the models (AIC_{min} = 721.32, χ^2 tests $p < 0.05$). Furthermore, interaction between therapist personae and client personae significant improved the descriptive power of these models (AIC_{min}(no-interaction) = 1137.18, AIC_{min}(interactions) = 721.32, $\chi^2(284) = 983.86$, $p < 0.001$). These results are in line with previous works suggesting that alliance is the product of the dyadic interaction between client and therapist [5]. We compare each one of the four character combinations produced by our method (see Sec. 4). We found that models trained with therapist character from client's text plus the client character from therapist's text achieved the best result out of any other possible combination of characters ($AIC(t\&c)$ = 721.32, AIC_{min}(others) = 781.08, $\chi^2(27) = 43.194$, $p < 0.05$). This character combination seems to capture two important aspects of alliance: the shared relationship bond, by considering characters across roles, and a therapist who narrates client-focused stories, instead of just recounting their own life-experiences. The best model achieved an AIC_{min} = 721.32 with K = 30 topics, P = 30 personae, and character cross interaction (client from therapist speech, and vice versa). No other model is within 10 AIC units from this result. Thus, our model selection procedure yields a single best model. This model is significantly better than the models with other choices of K and P ($\chi^2(2)$ = 119.09, $p < 0.001$).

Regression results.

The performance of regression models is presented in Table 2. Consistent with the results with LME models, using only therapist or client text performs poorly compared to using both therapist and client information. The poor performance from linguistic models supports that an individual's language use does not capture alliance information [2]. Including persona information about either the therapist or the patient significantly improved the performance over the separate therapist or patient linguistic models (t-test, $t(60) = 3.94$, $p < 0.001$ and t-test, $t(60) = 9.13$, $p < 0.001$ respectively). Furthermore, personae model also performs better than the supervised model that considers both therapist and client text (t-test, $t(60) = 7.01$, $p < 0.01$). However, no statistical differences ($p > 0.05$) were found between the model with interactions and the models with either the therapist or client's personae. The poor performance of the linguistic models suggests that the regression weights are over-fitting to specific instances of word usage. This further suggests that the personae model is successfully capturing higher-order interactions that are not represented in the vocabulary of the participants.

7. Personae Analysis

To understand what the Persona model captures from the data, we inspect the distributions of personae and topics. We focus on those clients and therapist that appear more than once to increase reliability of the analysis. Thus the subset of our data for this analysis contains n = 802 sessions, with 31 unique therapists and 204 unique clients. The number of sessions per therapist ranges from 2 to 105 ($\mu = 25.87$, $\sigma = 26.47$), for clients this ranged between 2 and 12 ($\mu = 3.93$, $\sigma = 2.19$).

Persona distribution.

Table 3 shows distribution statistics for the discovered personae. Interestingly, all 31 therapists changed their persona at least once. Thus, that our models are not just creating a persona for each therapist across all sessions, but instead grouping therapists into certain personae depending on the characteristics of a session. Moreover, this suggests that therapists are changing their personae in accordance to each client. In contrast, clients portrayed fewer personas on average. For most of the clients, therapists see between 2 and 3 personae (61.76%). This suggest that the client's personae are dynamic and change through the therapeutic process. Investigating if this change is meaningful and what it is capturing will be explored in a future work. With respect to the topics, clients describe the therapist's mode (i.e., most frequent persona) with topics related to *arguments*, *love*, and *court* (as in court of law but also "the ball is in your court"); therapists describe clients with the most popular persona using topics of *wellness*, *issues* and *expectations*, cognitive skills (such as *think* and *feel*), and religion-related terms (e.g., *god*, *religious*). With respect to their interactions, we observe 428 personae pairs out of the 900 possible ones. There was a significant difference in the alliance between personae pairs (Kruskal-Wallis, $H = 477.59$, $p < 0.05$). Once again, this result supports our claim that interactions between certain types of characters achieve higher levels of alliance than others.

Topic contribution.

We inspect the sign and magnitude of the Linear SVR coefficients as a measure of topic importance. The most important indicators of high alliance are clients that solve schedule conflicts (e.g., "should we schedule for next week?"), therapists that show empathy (e.g., "[no need] to be blaming yourself either, you worked hard on that paper", "I hope I'm not communicating judgment"), specially when clients are going through worrying or stressful moments (e.g., "it triggered a panic attack", "[I] worry too much about what i need to do tomorrow"). The top indicators of low alliance are uncommitted or distracted therapists (e.g., "let me check on that", "let me know next week"), therapists not completely understanding what the client is going through (e.g., "What's it like to just sit with this anxiety?", "I wonder what could have been, and for me") or invalidating a client's feelings (e.g., "I'd imagine that I would have made the decision to stay", "yeah yeah like just because you're feeling anxious doesn't mean you're going to [...]").

8. Conclusions

We presented a novel approach to understand alliance construction, using a notion of client and therapist personae. Personae are automatically discovered using an unsupervised approach—thus, not requiring additional manually labeled data. Our claim that alliance is captured by the interaction between the discovered character types is supported by: (1) LME models achieve significantly better results with interaction terms; (2) there is a significant difference between alliance ratings for certain pairs of therapist/clients, and (3) machine learning models with persona representations predict alliance significantly better than supervised linguistic baselines. There are two main limitations for this work: first, we did not make any effort in recovering downstream errors from the role matching and ASR systems, which might induce errors in the words selected for the topics, and, second, our assumption on how participants refer to each character does not consider the case of impersonal you, which might induce errors in the topics selected for the personae. Both of these limitations will be addressed in future work.

Acknowlegements

Special thanks to Jake Van Epps and the University of Utah Counceling Center for their contribution to this work. This work was supported by NIH, NSF and USC. NF is partially supported by the USC Annenberg Fellowship. VM is partially supported by Mexican Council of Science and Technology.

10. References

[1]. Morgan A, What is narrative therapy? An easy-to-read introduction, 1st ed. Dulwich Centre Publications, 2000.

[2]. Goldberg SB, Flemotomos N, Martinez VR, Tanana M, Kuo P, Pace BT, Villatte JL, Georgiou P, Epps JV, Imel ZE, Narayanan S, and Atkins DC, “Transforming psychotherapy research with machine learning: Alliance as a case in point,” submitted.

[3]. Lambert MJ and Barley DE, “Research summary on the therapeutic relationship and psychotherapy outcome.” Psychotherapy: Theory, research, practice, training, vol. 38, no. 4, 2001.

[4]. Thompson MN, Goldberg SB, and Nielsen SL, “Patient financial distress and treatment outcomes in naturalistic psychotherapy.” Journal of counseling psychology, vol. 65, no. 4, 2018.

[5]. Baldwin SA, Wampold BE, and Imel ZE, “Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance.” Journal of consulting and clinical psychology, vol. 75, no. 6, 2007.

[6]. Flückiger C, Del Re A, Wampold BE, and Horvath AO, “The alliance in adult psychotherapy: A meta-analytic synthesis.” Psychotherapy, 2018.

[7]. Taber BJ, Leibert TW, and Agaskar VR, “Relationships among client–therapist personality congruence, working alliance, and therapeutic outcome.” Psychotherapy, vol. 48, no. 4, 2011.

[8]. Miller WR, Moyers TB, Ernst D, and Amrhein P, “Manual for the motivational interviewing skill code (misc),” Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico, Tech. Rep, 2003.

[9]. Bamman D, O’Connor B, and Smith NA, “Learning latent personas of film characters,” in Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) Sofia, Bulgaria: Association for Computational Linguistics, Aug. 2013.

[10]. Jung CG, Two essays on analytical psychology. Routledge, 1943.

[11]. Kim H, Katerenchuk D, Billet D, Park H, and Li B, “Learning Joint Gaussian Representations for Movies, Actors, and Literary Characters,” in Proceedings of the Thirthy-Third AAAI Conference on Artificial Intelligence AAAI, 2019.

[12]. Finlayson MA, "A survey of corpora in computational and cognitive narrative science," International Journal for Language Data Processing, vol. 37, no. 1–2, 2013.

[13]. Riedl M, "Computational Narrative Intelligence: Past, Present, and Future," https://bit.ly/2rA3PTE, Oct. 2017, accessed: 2019-04-02.

[14]. Hall JM and Powell J, "Understanding the person through narrative," Nursing research and practice, vol. 2011, 2011.

[15]. Diers D, Speaking of Nursing–: Narratives of Practice, Research, Policy, and the Profession. Jones & Bartlett Learning, 2004.

[16]. Mishler EG, The discourse of medicine: Dialectics of medical interviews. Greenwood Publishing Group, 1984, vol. 3.

[17]. Nicolopoulou A and Trapp S, 17. Narrative interventions for children with language disorders: A review of practices and findings, 07 2018.

[18]. White M, White MK, Wijaya M, and Epston D, Narrative means to therapeutic ends WW Norton & Company, 1990.

[19]. Etchison M and Kleist DM, "Review of narrative therapy: Research and utility," The Family Journal, vol. 8, no. 1, 2000.

[20]. Vromans LP and Schweitzer RD, "Narrative therapy for adults with major depressive disorder: Improved symptom and interpersonal outcomes," Psychotherapy Research, vol. 21, no. 1, 2011.

[21]. Yao L, Peng N, Ralph W, Knight K, Zhao D, and Yan R, "Plan-and-write: Towards better automatic storytelling," in Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence. AAAI, 2019.

[22]. Hassanali K.-n., Liu Y, and Solorio T, "Using latent dirichlet allocation for child narrative analysis," in Proceedings of the 2013 Workshop on Biomedical Natural Language Processing. Sofia, Bulgaria: Association for Computational Linguistics, Aug. 2013.

[23]. Bamman D, Underwood T, and Smith NA, "A bayesian mixed effects model of literary character," in Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), vol. 1, 2014.

[24]. Coleman D, "Therapistclient fivefactor personality similarity: A brief report," Bulletin of the Menninger clinic, vol. 70, no. 3, 2006.

[25]. GROUP PMR, "Matching patients with alcohol disorders to treatments: Clinical implications from project match," Journal of Mental Health, vol. 7, no. 6, 1998.

[26]. Bower B, "Alcoholics synonymous: Heavy drinkers of all stripes may get comparable help from a variety of therapies," Science News, vol. 151, no. 4, 1997.

[27]. Peele S, "Ten radical things niaaa research shows about alcoholism," The Addictions Newsletter, vol. 5, no. 6, 1998.

[28]. Gibson J, Can D, Georgiou P, Atkins D, and Narayanan S, "Attention networks for modeling behavior in addiction counseling," in Proceedings of Interspeech, August 2017.

[29]. Can D, Marin R, Georgiou P, Imel ZE, Atkins D, and Narayanan SS, ""it sounds like...": A natural language processing approach to detecting counselor reflections in motivational interviewing," Journal of Counseling Psychology, 2015.

[30]. Howes C, Purver M, and McCabe R, "Investigating topic modelling for therapy dialogue analysis," in Proceedings of the IWCS 2013 Workshop on Computational Semantics in Clinical Text (CSCT 2013) Association for Computational Linguistics, 2013.

[31]. ——, "Linguistic indicators of severity and progress in online text-based therapy for depression," in Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality, 2014.

[32]. Horvath AO and Greenberg LS, "Development and validation of the working alliance inventory." Journal of counseling psychology, vol. 36, no. 2, 1989.

[33]. Hatcher RL and Gillaspy JA, "Development and validation of a revised short version of the working alliance inventory," Psychotherapy Research, vol. 16, no. 1, 2006.

[34]. Povey D, Ghoshal A, Boulianne G, Burget L, Glembek O, Goel N, Hannemann M, Motlicek P, Qian Y, Schwarz P, Silovsky J, Stemmer G, and Vesely K, "The kaldi speech recognition toolkit,"

in IEEE 2011 Workshop on Automatic Speech Recognition and Understanding. IEEE Signal Processing Society, Dec. 2011, iEEE Catalog No.: CFP11SRW-USB.

[35]. Sell G, Snyder D, McCree A, Garcia-Romero D, Villalba J, Maciejewski M, Manohar V, Dehak N, Povey D, Watanabe S et al. , “Diarization is hard: Some experiences and lessons learned for the jhu team in the inaugural dihard challenge,” in Proc. INTERSPEECH, 2018.

[36]. Peddinti V, Chen G, Manohar V, Ko T, Povey D, and Khudanpur S, “Jhu aspire system: Robust lvcsr with tdnns, ivector adaptation and rnn-lms,” in 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU) IEEE, 2015.

[37]. Flemotomos N, Martinez V, Gibson J, Atkins D, Creed T, and Narayanan S, “Language features for automated evaluation of cognitive behavior psychotherapy sessions,” Proc. Interspeech 2018, 2018.

[38]. Manning CD, Surdeanu M, Bauer J, Finkel J, Bethard SJ, and McClosky D, “The Stanford CoreNLP natural language processing toolkit,” in Association for Computational Linguistics (ACL) System Demonstrations, 2014.

[39]. Blei DM, Ng AY, and Jordan MI, “Latent dirichlet allocation,” Journal of machine Learning research, vol. 3, no. Jan, 2003.

[40]. Burnham KP and Anderson DR, “Multimodel inference: understanding aic and bic in model selection,” Sociological methods & research, vol. 33, no. 2, 2004.

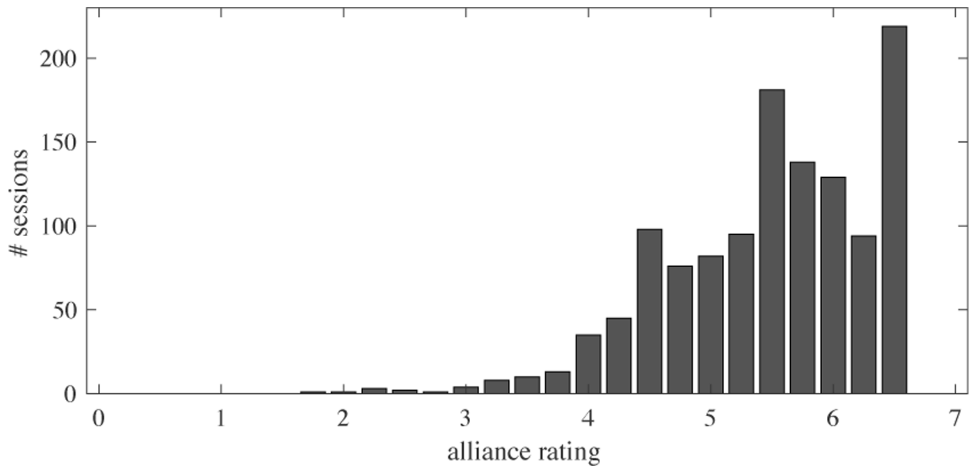

Figure 1:
Histogram of the therapeutic alliance ratings.

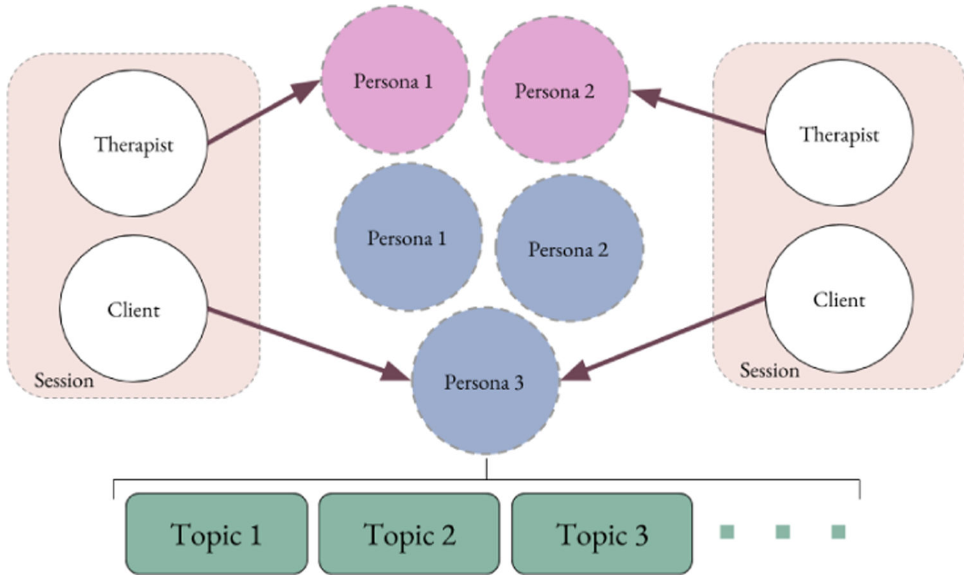

Figure 2:
Personae and topic distributions from psychotherapy sessions: Personae are distributions over the topics of conversation (shown in green). Sets are learned for clients and therapists independently (shown in blue and light pink respectively). Therapist and client get assigned to a single persona per session.

Table 1:

Number of sessions per client, per therapist and per (client, therapist) pair in the available dataset. Support is the total number of clients, therapists, or pairs.

	min	mean	max	support
# sessions per client	1	3.20	13	386
# sessions per therapist	1	30.88	131	40
# sessions per pair	1	3.11	13	397

Table 2:

Cross-validation estimation for mean (μ) and standard deviation (σ) of MSE for regression models (lower is better). Persona model performs significantly better than the baselines.

	Mean Squared Error	
	μ	**σ**
Therapist-only (unigram)	9.27	7.38
Therapist-only (bigram)	13.09	7.77
Client-only (unigram)	15.40	10.67
Client-only (bigram)	20.69	11.13
Therapist + Client (unigram)	3.04	1.62
Therapist + Client (bigram)	4.32	2.31
Personae (K = 30, P = 30)	**0.69**	0.51
-Client Only	**0.69**	**0.50**
-Therapist Only	0.70	0.54

Table 3:

Mean μ, standard deviation σ and mode for persona distribution per participant and joint distribution.

	μ	σ	**Mode (*n*)**
Therapist	10.25	6.61	29 (77)
Client	3.37	1.68	16 (50)
Client + Therapist	1.87	1.23	16 and 2 (8)