Curr Psychiatry Rep (2016) 18: 49
DOI 10.1007/s11920-016-0682-5

Computational Analysis and Simulation of Empathic Behaviors: a Survey of Empathy Modeling with Behavioral Signal Processing Framework

Bo Xiao [1] **· Zac E. Imel** [2] **· Panayiotis Georgiou** [1] **· David C. Atkins** [3] **·**
Shrikanth S. Narayanan [1]

Published online: 28 March 2016

Abstract Empathy is an important psychological process that facilitates human communication and interaction. Enhancement of empathy has profound significance in a range of applications. In this paper, we review emerging directions of research on computational analysis of empathy expression and perception as well as empathic interactions, including their simulation. We summarize the work on empathic expression analysis by the targeted signal modalities (e.g., text, audio, and facial expressions). We categorize empathy simulation studies into theory-based emotion space modeling or application-driven user and context modeling. We summarize challenges in computational study of empathy including conceptual framing and understanding of empathy, data availability, appropriate use and validation of machine learning techniques, and behavior signal processing. Finally, we propose a unified view of empathy computation and offer a series of open problems for future research.

Keywords Empathy · Computational modeling · Analysis · Simulation · Behavioral signal processing

This article is part of the Topical Collection on *Psychiatry in the Digital Age*

✉ Shrikanth S. Narayanan
shri@sipi.usc.edu

Bo Xiao
boxiao@usc.edu

Zac E. Imel
zac.imel@utah.edu

Panayiotis Georgiou
georgiou@sipi.usc.edu

David C. Atkins
datkins@u.washington.edu

[1] SAIL, Department of Electrical Engineering, University of Southern California, Los Angeles, CA, USA

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

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

Introduction

Definition of Empathy

The word *empathy* has its origins in the Greek word εμπάθεια, meaning an inward aspect of "I feel" or "I suffer." Its usage in the psychology literature started in 1909 with Titchener's translation of the German term "*Einfühlung*" [1].

The term of empathy takes multiple interpretations. Hoffman defined it as "an affective response more appropriate to another's situation than one's own" [2], while Batson listed eight distinct phenomena that are all named empathy [3]. The discussion of empathy's definition continues in a recent summary by Cuff et al. [4]. Despite conceptual variations, consensus on the understanding of empathy consists of three major subprocesses [3, 5, 6], including

- Emotional simulation - An affective response which often entails sharing the emotional state
- Perspective taking - A cognitive capacity of knowing another's internal states including thoughts and feelings
- Emotion regulation - Regulating personal distress from the other's pain to allow compassion and helping behavior.

Interdisciplinary research on empathy modeling has broadened and deepened the understanding of empathy. Preston suggested that a perception-action model has the explanatory power to integrate different views of empathy into a common mechanism framework. The model states that "attended perception of the object's state automatically activates the subject's representations of the state, situation, and object and the activation of these representations automatically primes or generates the associated autonomic and somatic responses, unless inhibited" [7]. Decety and Jackson modeled empathy as "parallel and distributed processing in a number of dissociable computational mechanisms," including shared neural representations, self-awareness, mental flexibility, and emotion regulation, which are supported by specific neural systems [5]. De Vignemont and Singer argued that empathic brain response may be contextual rather than automatic, modulated by the appraisal processes, taking into account factors such as information about the emotional stimuli, their situational context, characteristics of the empathizer, and his/her relationship with the target [8].

Importance of Empathy

Acquired during evolution [7, 9], empathy likely serves to motivate sympathetic, helping, cooperative, and prosocial behaviors and facilitates social communication [6, 8]. In the context of psychotherapy, Elliott et al. have conducted a meta-analysis that revealed an overall positive correlation of 0.31 between therapist empathy and client outcome. Thus empathy is among the most consistent predictors of psychotherapy outcome available [6].

In clinical fields of oncology and general medical practice, positive correlations between empathy measures and patient outcomes have also been found in meta-analyses [10, 11]. Moyers and Miller also summarized the importance of empathy in psychotherapy and proposed that empathic listening skills should be emphasized in hiring and training therapists [12]. Concerning whether empathy may be taught, a recent review concluded that empathy training tends to be effective in general [13].

Challenges

There are still important challenges in promoting empathy in clinical settings. Empathy is in part an internal mental process, which is difficult to gauge directly by observation. Measurement of it relies on human perception and subjective assessment, either by the client, the therapist, or an outside reviewer [6]. These measures vary from the true psychological process, thus being fundamentally a probabilistic estimate with associated statistical inaccuracy. They may also be biased, exacerbating the problem of coder reliability. Human ratings also tend to be time consuming and hence is prohibitive for large-scale measurement of therapist empathy [14•]. The gain of empathy from training may decay over time, while day-to-day monitoring and reinforcement of empathy by human experts is generally out of reach. In addition to being relatively slow, human ratings may not be sufficiently sensitive to capture particular nuanced and latent facets of the empathic process (e.g., synchrony). As a result, research on how to decode human behaviors with respect to empathy expression, perception, and action is still in its early stage, partly due to physical constrains on acquiring large amounts of data of therapist behaviors against empathy evaluations.

Empathy and Computation

Computational methods provide potential solutions to the aforementioned problems with scale and specificity. Recent technological advances have enabled easy, large-scale, and widely deployable audio, visual, and physiological sensing abilities; concurrent advances in signal processing and machine learning techniques have made possible for computers to analyze complex human behaviors from vast amounts of diverse multimodal data. If automated computational methods are able to discern empathy, the advantages are clear; machines provide objective assessments and enable unconstrained sensing and computational bandwidth to support scalability.

In this paper, we conduct a survey on computational topics related to empathy: (i) analysis of empathic human behaviors, through multimodal observation signals, in "Empathy Analysis" section, and (ii) simulation of empathic human behaviors, through design of artificial computer agents, in "Empathy Simulation" section.

In "Challenges and Future Directions" section, we discuss key issues faced in empathy computation and propose future research directions. We conclude in "Conclusion" section.

Empathy Analysis

In behavioral studies of empathy, typically human raters (who are often external to the interaction/data generation setting) use behavioral cues of the target to infer and annotate whether a particular empathic process has occurred (e.g., a group of behavioral cues proposed by Riess [15] and an analysis of the contribution of different cues by Regenbogen et al. [16]). Likewise, computational empathy analysis studies how to capture and model multimodal behavioral cues for detecting empathy.

Two kinds of research methodologies are commonly applied.

- Feature analysis - Finding behavioral cues that correlate with human annotator-derived empathy ratings through statistical analyses, a common method in behavioral sciences.

- Prediction - Data driven computational learning of models (using machine learning techniques) that serve as functions mapping automatically measured behavioral cues to empathy ratings. The performance of the automated prediction is typically evaluated by comparing machine assessments against human expert ratings on new or held-out interactions not seen in model construction [17].

The standard in clinical psychology and psychiatry is to build and evaluate models in a complete dataset (e.g., to fit a regression model with various correlates of empathy). In engineering approaches, *prediction* is a much stronger test than *correlation*. It partitions data into mutually exclusive training and evaluation sets to establish validity and generalizability of results. In the following, we describe both types of studies, but the readers should note that prediction refers to the situation when a new "test" set is used and is generally a more rigorous test of a particular hypothesis.

As an emerging field, computational empathy analysis has been pursued most notably in two domains. Firstly, in addiction counseling using motivational interviewing (MI) [18], empathy is a key index for treatment fidelity [19]. Human experts use the *Motivational Interviewing Treatment Integrity* (MITI) manual [20] to code the degree of therapist empathy in an interaction on a Likert scale. MITI defines empathy as "the extent to which the clinician understands or makes an effort to grasp the client's perspective and feelings," emphasizing the cognitive component of empathy.

Secondly, in four-person casual conversations, the researchers operationally defined empathy as emotion contagion [21], emphasizing the affective component of empathy. Human coders marked the empathy states of each pair of interlocutors on the time line.

Though in its early stage, computational empathy analysis has examined a number of multimodal behavioral cues. In addition, *entrainment* (synchrony)—an interaction process wherein behaviors of interlocutors becoming more similar or coordinated—is a phenomenon that is tied closely to empathy, based on the theory of perception-action link and the function of mirror neurons [7, 9, 22]. Modeling entrainment across various modalities serves as an indirect but useful mechanism for quantifying empathy.

Lexical Cues

Spoken language encodes a multitude of information including a speaker's intent; emotions; desires; and other physical, cognitive, and mental state and traits (e.g., speaker age and gender). By analyzing the language transcripts of interactions, we may infer the empathy processes that are driving, and reflected in, the language expressions (e.g., qualitative findings on empathic word use by Coulehan et al. [23]).

Xiao et al. have used N-gram language models (see Table 1) [24] of empathic vs. other (background) utterances of the therapists in MI-type counseling [25•]. They showed that a maximum likelihood classifier (see Table 1) based on these language models was useful to automatically identify empathic utterances. Further, utterance level evidences of empathy can be summed to derive measures that can better *correlate* with interaction session level empathy ratings (i.e., MITI codes).

Extending this work, Chakravarthula et al. proposed a model that considers the therapist's likelihood to transition among high- vs. low-empathy states over time using a hidden Markov model (see Table 1) [26], instead of assuming a static state of empathy throughout the interaction [27•]. They showed that the dynamic model provided improved *predictions* of the session-level assessments offered by human experts compared to the static model while providing short-term empathy information.

The above N-gram language model-based methods do not exploit the semantic meaning of words. Linguistic features such as those generated by the *Linguistic Inquiry and Word Count* (LIWC) software [28] associate words with categories of various psychological processes, personal concerns, spoken categories, etc. Moreover, novel computational methods afford affective text analyses to be applied broadly beyond words specified in the lexica [29]. Computational psycholinguistic norms (PNs; see Table 1) [29] further expand the ability to include both affect states and word's relation to additional cognitive processes (e.g., age of acquisition, imageability, and gender ladenness). Gibson et al. compared LIWC and PN features to N-gram features in predicting therapist empathy ratings, showing that although N-gram features performed the best, LIWC and PN features provided complementary information resulting in boosted *prediction* performance by feature fusion [30••].

The above methods investigate language cues that directly correlate with and can predict empathy. Although these cues appear to be effective, their ties to psychological theories about empathy largely remain implicit. On the other hand, analysis of language style synchrony investigates one possible realization of the perception-action link. Lord et al. extracted LIWC features on each speaking turn of the therapist/client and quantified if the same category of words appeared both in the therapist's turn and the client's turn [31•]. As a result, they found 11 word categories that associated with stronger synchrony in high-empathy sessions. Language style synchrony has even stronger correlation to empathy than the well-accepted traditional indicator—count of *reflections* by the therapist.

Table 1 Explanation of technical terms

N-gram model	A sequence of N contiguous words is named an N-gram. Probability of a word sequence can be described by the probabilities of N-grams, e.g., P (it sounds like) equals the product of P(it), a uni-gram; P (sounds|it), a bi-gram; and P (like|it sounds), a tri-gram. We may assume any word only depends on the previous two words, so that a tri-gram model can derive the probability for a word sequence of arbitrary length.
Maximum likelihood classifier	Denote likelihoods derived by competing classes as $P(x\|C)$, where C is the class label and x is an observation. Classify x to class $C^* = \arg\max_c P(x\|C)$.
Hidden Markov model	A statistical model composed by a sequence of unobserved (hidden) nodes and observed output attached to each hidden node. Hidden nodes have discrete states depending only on the previous node (i.e., Markovian). A state transition probability matrix, a conditional output probability, and an initial state distribution compose the statistical characteristics of the model.
Psycholinguistic norms	Indices in range −1 to 1, derived based on manual annotation on a small set and automatic estimation for any word using semantic similarity. For example, "love" and "suicide" have valence scores 0.93 and −0.94, respectively.
Pitch	In auditory terms, the relative level of tone perceived by the ear, which depends on the count of vibrations per second by the vocal folds. In acoustic terms, estimated as the fundamental frequency of the speech signal in the unit of hertz.
Energy	Logarithm of mean-squared value of speech signal, an estimate of speech intensity in acoustic terms, and loudness in auditory terms.
Jitter	Estimate of the variation of fundamental period, calculated as the average time difference of pitch reciprocals.
Shimmer	Estimate of the variation of speech intensity, calculated as the average time difference of speech energy.
MFCC	Coefficients derived through discrete cosine transformation of a log power spectrum on a mel-scale of frequency. The mel-scale approximates the non-linear frequency bands in human auditory system.
PCA	An orthogonal transformation on a vector of variables, resulting linearly uncorrelated variables named principal components, which are listed in the order of variance in the observed data.
KLD	A non-symmetric measure of the difference between two probability distributions. For example, $D\ (P\|\|Q)$ denotes the information loss when a distribution Q is used to approximate P, defined as e.g., $D(P\|\|Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$ for discrete distributions.
Dynamic Bayesian network	A probabilistic graphic model composed by a set of nodes and edges as a directional acyclic graph. Each node represents a random variable, while an edge connecting two nodes represents conditional dependency between them. Given some nodes observed, there exist efficient algorithms to derive the posteriors of other nodes in the graph. The dynamic aspect denotes a network structure that repeats along time.
Formal language	A set of strings of symbols that are constrained by specific rules, e.g., grammar and logical operations.
Naive Bayes model	A family of probabilistic classifiers based on Bayes' theorem with strong independence assumptions between the features. Following the *maximum a posteriori* decision rule, the class label is derived as $y = \arg\max_{k\in\{1,\cdots,K\}} p(C_k)\prod_{i=1}^{n} p(x_i\|C_k)$.
Decision tree	A model representing an algorithm, where branching operations take place at nodes through certain comparison functions. Final decisions are made at the leaf nodes.
Reinforcement learning	An approach concerning how an agent takes actions in an environment so as to maximize some notion of cumulative reward. It balances two aspects in an online learning process, exploration of unseen territory and exploitation of current knowledge.
Support vector machine	A type of binary classifier, having a property that the dividing hyper-plane of the two classes are furthest to any sample of the two classes in the training set, so called "large margin" property. The dividing hyper-plane can be linear or non-linear using the kernel method.

Vocal Cues

Human vocal expression is highly dependent on internal state, and as such, it is linked to empathy. This has been supported by diverse work; e.g., brain areas important for prosodic mechanisms are linked to empathic ability [32], and empirically, prosodic continuity (e.g., therapist continued the intonation/rhythm of the client's preceding turn) by the therapist has been associated with higher empathy [33].

Xiao et al. studied whether prosodic patterns related to empathy assessments [34••]. They extracted prosodic features for each speech segment of the therapist and the client, including vocal pitch, energy, jitter, shimmer, and speech segment duration (see Table 1). Joint distributions of these features were examined for *correlation* with empathy. The results suggested a group of significant empathy indicators, which were able to *predict* high vs. low empathy. For example, increased distribution of medium-length segment with both high energy and high pitch associated with lower-empathy assessments. This finding suggests that raised intonation and louder voice by the therapist may be perceived as signaling lower empathy.

Further to direct vocal cues, interlocutor vocal entrainment serves as an indirect feature for empathy. Imel et al. investigated vocal entrainment through the correlation of mean fundamental frequencies (pitch) [35] between interacting therapist and standardized patient (SP) [36•]. They found strong correlation (0.71) that did not exist in fake interactions with random pairings of therapists and SPs. Moreover, this correlation was higher in high-empathy sessions compared to low-empathy ones, demonstrating the link between entrainment and empathy.

Xiao et al. modeled entrainment with a more detailed measure of acoustic similarity [37•]. They extracted MFCCs (i.e., mel frequency cepstrum coefficients; see Table 1) [35] and pitch features from the speech of interacting therapists and SPs. These features defined the principal component analysis (PCA; see Table 1) [38] spaces of the therapist/SP. Kullback-Leiber divergence (KLD; see Table 1) [39] was employed to compute the similarity of PCA components. They found significant *correlation* between statistics of turn-level KLDs and human-specified empathy ratings.

Xiao et al. investigated speech rate (i.e., number of words, syllables, or phonemes in a unit of time) entrainment and its link to empathy [40••]. They showed that the mean absolute difference of speech rates between the therapist and the client *correlated* with therapist empathy. In addition, statistics of speech and silence durations were also significant *correlates* of empathy. These features provided complementary information to the prosodic features in [34••] in *predicting* sessions assessed as high vs. low empathy. The above three studies lend support to the perception-action model of empathy from vocal cues.

Facial Expression and Reaction Timing Cues

Facial expressions also carry rich emotional information [41]. Kumano et al. investigated if the co-occurrence of facial expression patterns among the interlocutors could *predict* the empathy labels [42]. They discretized facial expressions into six types and modeled empathy state in three classes as *empathy*, *unconcern*, and *antipathy*. A dynamic Bayesian network model (see Table 1) [43] was constructed to associate empathy states with facial expressions and gaze directions along time. Experiment results showed that facial expressions were effective predictors of empathy labels.

Kumano et al. extended this framework by investigating reaction timing and facial expression congruence information [44•]. They demonstrated that these two aspects were related to the annotated empathy labels (e.g., a congruent but delayed reaction in facial expression is less likely to have an empathy label). By further incorporating annotations of head gesture types, they improved the accuracy of empathy state *prediction*.

Moreover, Kumano et al. studied the inference of empathy labels by multiple human annotators [45•]. Instead of assigning one class label for empathy, they estimated the distribution of empathy labels by a group of evaluators. They found that training the model with multiple annotations outperformed training with only the majority-voted empathy labels.

Empathy Simulation

Empathy simulation aims at the dual problem to empathy analysis, i.e., artificial embodiment and display of empathic behaviors in virtual or robotic agents, which are perceived by human users. So far, it is still impossible to recreate the human neural cognitive system in machines, so that “truly empathic” avatars are impossible to make. However, a simulation of human-like behavior that invokes a perception of empathy by the user is feasible and useful for experimentation and applications [46]. The methodology usually includes a theory- or practice-inspired design of an “empathy-embedded” artificial system and human evaluation of its effectiveness. Work in this field can be roughly summarized in two directions—driven by a computational model of the emotion space that is inspired by theory or driven by user and context modeling in specific applications. The former attempts to simulate the empathy process in human brain, expecting such design to influence the behaviors of computational agents to become empathic, while the latter tracks user’s emotional state and context in the application and reacts with appropriate pre-defined expressions that can be perceived as empathy.

Computational Model of Emotion Space

The *emotion contagion* phenomenon, as one element of empathy, has been a relatively simpler target of empathy simulation. Riek and Robinson conducted a preliminary study to test the empathy effect of facial expression mimicry by a robot [47]. They found that facial expression mimicry—as a way to mimic emotion—helped in increasing the satisfaction of human users. Gonsior et al. investigated mimicking users’ facial expressions with a talking robot [48]. They found that users rated the robot mimicking facial expressions as being more empathic than the one showing a neutral face.

However, emotion mimicry may not be the entire characteristic of empathy, and the intensity may be modulated by other factors as De Vignemont suggested [8]. A study by Becker et al. found that in a scenario of human-machine card game, emotion mimicry by the virtual agent in a constant manner increased the stress of the user [49]. Thus, parameterization of the emotion space and a model of behavior modulation become critical in empathy simulation.

In light of this, Boukricha et al. proposed a scheme of three-dimensional emotion space including pleasure, arousal, and dominance (PAD) [50, 51•]. In addition, a three-step model was proposed to produce an empathic reaction: (i) *empathy mechanism*—an internal imitation of perceived facial expressions and an emotional feedback that represents the perceived emotion; (ii) *empathy modulation*—modulation of empathic emotion (i.e., an emotion likely invoking perceived empathy by human users) as an interpolation of the perceived and own emotion (mood) states in the PAD space, weighted by degrees of factors such as *liking* and *familiarity*; and (iii) *expression of empathy*—the modulated emotion states triggering facial, vocal, and verbal expressions accordingly. The authors found in experiment that the degree of empathy expressed by a virtual agent is consistent with the tuning parameter of *liking*.

Ochs et al. proposed a formal language (see Table 1) based model for simulating empathic emotion [52•]. Firstly, *belief*, *uncertainty*, and *intension* were defined as notions of mental states of the virtual agent driven by the dialogue situations. Secondly, several types of emotions (e.g., *satisfaction*) were defined in formal logic based on these notions. Type, intensity, target, triggering event, and the affected intension of the emotion state of an agent (real or virtual) were incorporated in the model. Further, empathic emotion was elicited when an agent believed that another agent had a certain type of emotion. To ensure that a virtual agent is "well intentioned," it works under the axiom that users have neither negative nor lacking positive emotions.

Rodrigues et al. proposed an empathic emotion simulation model [53, 54••] that was in line of the previous work. They denoted an event with four elements including *subject*, *action*, *target*, and *parameters*. An emotion state was then denoted also as four elements including *type*, *valence*, *intensity*, and *cause* (event causing the emotion). The first step in the empathy process model was a scheme appraising the events being the causes of the other's emotion toward oneself, i.e., "putting oneself in another's shoes," resulting an elicited emotion. In addition, there was an emotion recognition scheme via observed cues, resulting in a set of recognized emotions. A potential empathic emotion was selected from the elicited and recognized emotions, which was then modulated by a group of factors including *mood*, *similarity*, *personality*, and *affective link*, following the theory by De Vignemont [8]. The modulated emotion was finally expressed through reactive behaviors. In the experiment, human evaluators assessed the virtual agent with such model as having more prosocial characteristics in several aspects. All these simulation models hold the promise both as a research tool to explore specific hypotheses about empathic processes and in implementing useful human-machine interface applications.

Application-Oriented User and Context Modeling

Data-driven approaches for empathy simulation learn the context of human empathic behavior exemplars, i.e., modeling when to display which expression. For example, McQuiggan and Lester designed the *CARE* framework [55, 56]. They collected the behaviors of a virtual agent that was manipulated by a human acting in an empathic manner (e.g., feeling frustration when the user is losing the game). The recorded data were used to train naive Bayes and decision tree models [38] (see Table 1), which determined both when and how the virtual agent should mimic human empathic expressions. Human evaluation showed that there was no significant difference of judged appropriateness between the model-generated and human-manipulated behaviors in the application.

User state and context modeling may also facilitate a proper reaction strategy. Leite et al. designed a chess game companion robot named iCat [57•]. They tracked a user's emotional valence in positive and negative states based on gaze and facial expression of the user and the context of winning or losing in the game. A set of empathic expressions by the robot were prepared and applied either randomly or adaptively according to a reinforcement learning algorithm (see Table 1) [58] to maximize the probability of user's positive emotion. Children playing with the robot rated the empathic version with higher engagement, helping, and self-validation, compared to a neutral control. However, among the two empathic versions, the adaptively reacting one did not outperform the random one, possibly due to short interaction time to learn an optimally customized strategy.

Leite et al. carried out another study where iCat accompanied two human players in a chess game [59•]. During the game, iCat commented to one player with empathic expressions (e.g., feeling sad when the user is losing) while being neutral to the other. User's situations in the game (winning or losing and good or bad move) were used to estimate the affective state of the user and to determine the corresponding reaction of the robot. Facial expressions and verbal comments of the robot were employed as means of expression. As a result, the player to whom the robot reacted empathically rated the robot higher on companionship, reliable alliance, and self-validation.

D'mello et al. built a pedagogical virtual agent named *affective AutoTutor*, which acted in an empathic and motivational manner toward students [60•]. The system prepared in advance a set of facial, prosodic, and verbal responses of the AutoTutor that may be empathic, e.g., saying "I know this material can be difficult, but I think you can do it" for addressing a frustrated student. It detected user's conversational cues, facial expressions, and body postures, which were all integrated to derive the estimate of user's emotional state. A rule-based scheme was developed to select the proper response. Experiments showed that students with low prior knowledge in the subject gained more from the affective AutoTutor compared to a neutral version.

Challenges and Future Directions

In this section, we review some of the challenges that remain and offer possible future directions.

The Loaded Concept of Empathy

One of the main issues with empathy computation is that it is a complex term with task-specific significance and interpretation. Our primary focus here is on the cognitive perspective in addiction counseling and on emotion contagion in social interactions. The task-specific studies and the variations among the target-domain empathy interpretations may limit knowledge transfer.

Empathy is a complex construct that is conveyed through multimodal behavioral cues and involves two or more entities in communication. Even a single empathy process has to bring together at least behavior stimuli, behavior perception, empathic resonation, and empathic expression [61]. Researchers have to acknowledge the complexity of empathy and carefully position their work with respect to the definition and context of target empathic behaviors.

Data and Analysis Techniques

In empathy analysis, data is currently the primary limiting factor in both quantity and variety. Existing works have pulled audio recordings from a few large-scale psychotherapy studies totaling to thousands of sessions [62–67]; however, only a small fraction was finely annotated—in terms of both psychological assessments of mental and behavioral states, and having time-marked transcripts to train and validate automatic speech- and language-processing systems. The work by Kumano et al. has employed a small dataset totaling to a few hours [21, 42, 44•, 45•].

The variety of data is limited with respect to modalities and scenarios. In available psychotherapy data, only audio is typically recorded while video and physiology data are not collected. There are many domains such as education, customer service, and medical care that covet empathic interactions. Though studied extensively in their respective fields, these interactions have not been analyzed in terms of empathy quantification.

Machine learning methods have been widely used in predicting empathy annotations, including linear regression [38], support vector machines (see Table 1) [68], and dynamic Bayesian network [43]. The limited size of annotated data samples has constrained the application of more advanced learning techniques due to overfitting—model tuned to specific small dataset but not able to generalize for new data [17]. Careful data split in cross validation and appropriate application of learning algorithms are vital to gaining solid conclusions and avoiding pitfalls [69].

Manual annotations of behavioral cues have been needed for empathy analysis in varying degrees. Automation and integration of behavioral signal acquisition, processing, and assessment within a unified system are the limiting factors toward large-scale implementation of empathy analysis. In view of this, the authors of this paper are developing a pipeline of speech- and language-processing modules for the end-to-end analysis of empathy (among other targets) in addiction counseling and general human dyadic interactions in the future [70]. The system links modules such as voice activity detection (VAD; establishing where speech occurs [71]), diarization (determining who is speaking and when [72, 73]), automatic speech recognition (ASR; transcribing what is said [74, 75]), speaker role identification (e.g., as therapist or client), and empathy detection from the therapist's spoken words and acoustic cues.

A Unified View Under Behavioral Signal Processing

Techniques of empathy simulation provide a platform to test characteristics of empathic behavior. Applications of empathy-embedded virtual reality and human companion robots are growing with potential contributions to mental health care [76–79]. However, there is a gap between theory-based empathy simulation and application-oriented, hand-crafted empathic behaviors. Moreover, knowledge gained from empathy analysis has not been fully transferred to the design of empathic expressions in the simulation.

We propose a unified view for empathy computation following the method of behavioral signal processing (BSP) [80••], as shown in Fig. 1. Each interactant's expression and perception are critical in the communication, mediated by behavioral signals. Besides the interactants, many settings involve an observer or evaluator (e.g., a trainer of therapist), that is outside the interaction, whose characteristics should also be modeled. A

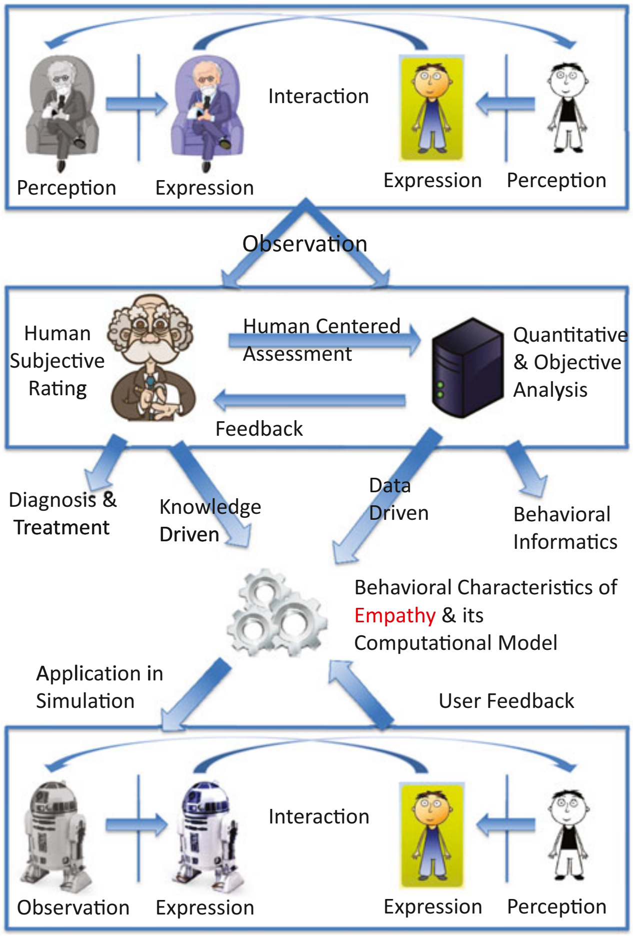

Fig. 1 A unified view of empathy analysis and simulation under BSP framework

computational model of empathy and understanding of its behavioral characteristics are central to both analysis and simulation. Such a model may be learned in analysis, applied in simulation, and further refined based on user feedback.

Based on this view, we list a number of open research questions related to empathy computation.

- Behavioral cues of empathy - Behaviors that express empathic states or that are perceived as being empathic by real humans. What are the cues that human express and perceive? Do they depend on the human's own state, the interlocutor's state, their mutual influence, and any relevant context? How do cues interplay across multiple modalities?
- Features - Behavioral signals that are derived from measurement, analysis, and modeling of behavior observations. What are the optimal ways to extract useful features for discerning one's empathy state? How to cope with individual disparity in measurement? How to fuse multiple features?
- Expression, perception, and evaluation - How to detect or manage the iterative process of expression and perception, particularly with respect to catching an *empathic opportunity* after an expression that entails empathic reaction? How to adapt to individual's subjective perception that weights various cues unevenly?
- Dynamics - How to track or manage the interaction along time with respect to modeling the evolvement of each individual's empathy state? How to derive an overall impression of empathy based on momentary assessments given the dynamics?
- Knowledge transfer - What are the computational structures and relations that are in common for empathy in varying scenarios? How to adapt empathy models to domain-specific representations?

Conclusion

Empathy is an important and complex neuro-cognitive process and serves a significant function in human interaction. It is multifaceted in its conceptual interpretation and application. Quantification of empathy and increased empathic behaviors can have a profound impact in a wide range of human-centered applications. Computational empathy analysis and simulation are emerging and encouraging new research directions, and we have attempted to summarize some of these in this paper.

Empathy analyses using multimodal signal processing and machine learning methods have identified useful features and models for empathy prediction. Empathy simulations have employed theory-based empathy elicitation mechanisms through emotion modeling and user context modeling-based empathy embedding in specific applications. Nevertheless, challenges remain in task-specific interpretations of empathy, in data sparsity, automatic behavior processing, and knowledge transfer between analysis and simulation of empathy. We have proposed a unified view of empathy modeling under the BSP framework and listed a series of open problems for the future. We believe the synergistic efforts in psychology, psychiatry, signal processing, machine learning, robotics, and artificial intelligence would facilitate gaining a deeper understanding of empathy and create new possibilities for empathy promotion via computational means.

Compliance with Ethical Standards

Conflict of Interest Bo Xiao, Zac E. Imel, Panayiotis Georgiou, and Shrikanth S. Narayanan declare that they have no conflict of interest.
David C. Atkins reports grants from NIAAA and NIDA.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1. Titchener EB. Lectures on the experimental psychology of the thought-processes. Macmillan; 1909.
2. Hoffman ML. Empathy and moral development: implications for caring and justice. Cambridge University Press; 2001.
3. Batson CD. In: Decety J, editor. These things called empathy: eight related but distinct phenomena. CRP press; 2009. p. 16–28.
4. Cuff BM, Brown SJ, Taylor L, Howat DJ. Empathy: a review of the concept. Emot Rev. 2014.
5. Decety J, Jackson P. The functional architecture of human empathy. Behav Cogn Neurosci Rev. 2004;3(2):71–100.
6. Elliott R, Bohart AC, Watson JC, Greenberg LS. Empathy. Psychotherapy. 2011;48(1):43–9.
7. Preston SD, De Waal F. Empathy: its ultimate and proximate bases. Behav Brain Sci. 2002;25(01):1–20.
8. De Vignemont F, Singer T. The empathic brain: how, when and why? Trends Cogn Sci. 2006;10(10):435–41.
9. Iacoboni M. Imitation, empathy, and mirror neurons. Annu Rev Psychol. 2009;60:653–70.
10. Derksen F, Bensing J, Lagro-Janssen A. Effectiveness of empathy in general practice: a systematic review. Br J Gen Pract. 2013;63(606):e76–84.
11. Lelorain S, Brédart A, Dolbeault S, Sultan S. A systematic review of the associations between empathy measures and patient outcomes in cancer care. Psychooncology. 2012;21(12):1255–64.
12. Moyers TB, Miller WR. Is low therapist empathy toxic? Psychol Addict Behav. 2013;27(3):878.
13. van Berkhout ET, Malouff JM. The efficacy of empathy training: a meta-analysis of randomized controlled trials. J Couns Psychol. 2015.
14.• Atkins DC, Steyvers M, Imel ZE, Smyth P. Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. Implement Sci. 2014;9(1):49. **This work used topic models to analyze provider fidelity based on text of the psychotherapy interaction, illuminating the possibility of scaling up the evaluation through computational means.**
15. Riess H. Biomarkers in the psychotherapeutic relationship: the role of physiology, neurobiology, and biological correlates of E.M.P.A.T.H.Y. Harv Rev Psychiatry. 2011;19(3):162–74.
16. Regenbogen C, Schneider DA, Finkelmeyer A, Kohn N, Derntl B, Kellermann T, et al. The differential contribution of facial expressions, prosody, and speech content to empathy. Cognition & emotion. 2012;26(6):995–1014.
17. Duda RO, Hart PE, Stork DG. Pattern classification. Wiley; 2012.
18. Miller WR, Rollnick S. Motivational interviewing: helping people change. Guilford Press; 2012.
19. Miller WR, Rose GS. Toward a theory of motivational interviewing. Am Psychol. 2009;64(6):527.
20. Moyers, T., Martin, T., Manuel, J., Miller, W., Ernst, D.: Revised global scales: Motivational Interviewing Treatment Integrity 3.0 (2007)
21. Kumano, S., Otsuka, K., Matsuda, M., Yamato, J.: Analyzing perceived empathy/antipathy based on reaction time in behavioral coordination. In: Automatic Face and Gesture Recognition, pp. 1–8. IEEE (2013)
22. Delaherche E, Chetouani M, Mahdhaoui A, Saint-Georges C, Viaux S, Cohen D. Interpersonal synchrony: a survey of evaluation methods across disciplines. IEEE Trans Affect Comput. 2012;3(3): 349–65.
23. Coulehan JL, Platt FW, Egener B, Frankel R, Lin CT, Lown B, et al. “Let me see if I have this right”: words that help build empathy. Ann Intern Med. 2001;135(3):221–7.
24. Kneser R, Ney H. Improved backing-off for m-gram language modeling. In: International conference on acoustics, speech, and signal processing, vol. 1. IEEE; 1995. p. 181–4.
25.• Xiao B, Can D, Georgiou PG, Atkins DC, Narayanan SS. Analyzing the language of therapist empathy in motivational interview based psychotherapy. Proc. APSIPA ASC; 2012. **This work was the first to find empathy correlates based on N-gram language models**.
26. Rabiner LR, Juang BH. An introduction to hidden Markov models. IEEE ASSP Mag. 1986;3(1):4–16.
27.• Chakravarthula SN, Xiao B, Imel ZE, Atkins DC, Georgiou P. Assessing empathy using static and dynamic behavior models based on therapist’s language in addiction counseling. Dresden: Proc. Interspeech; 2015. **This work improved text based empathy analysis through modeling the empathy states of the therapist in a dynamic fashion changing along time.**
28. Pennebaker, J.W., Booth, R.J., Francis, M.E.: Linguistic Inquiry and Word Count (LIWC) (2007). URL http://www.liwc.net/
29. Malandrakis N, Potamianos A, Iosif E, Narayanan S. Distributional semantic models for affective text analysis. IEEE Trans Audio Speech Lang Process. 2013;21(11):2379–92.
30.•• Gibson J, Malandrakis N, Romero F, Atkins DC, Narayanan S. Predicting therapist empathy in motivational interviews using language features inspired by psycholinguistic norms. Dresden: Proc. Interspeech; 2015. **This work extended lexical analysis of empathy to include linguistic and paralinguistic features.**
31.• Lord SP, Sheng E, Imel ZE, Baer J, Atkins DC. More than reflections: empathy in motivational interviewing includes language style synchrony between therapist and client. Behav Ther. 2015;46(3): 296–303. **This work investigated the entrainment aspect of language style in order to model empathy from the spoken language.**
32. Aziz-Zadeh L, Sheng T, Gheytanchi A. Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. PLoS One. 2010;5(1):1–8.
33. Weiste E, Peräkylä A. Prosody and empathic communication in psychotherapy interaction. Psychother Res. 2014;1–15.
34.•• Xiao B, Bone D, Van Segbroeck M, Imel ZE, Atkins D, Georgiou P, et al. Modeling therapist empathy through prosody in drug addiction counseling. Proc. Interspeech; 2014. **This work investigated prosodic correlates of empathy through quantizing the prosodic cues and finding the joint distribution of these cues**.
35. Rabiner L, Schafer R. Theory and applications of digital speech processing. 1st ed. Prentice hall; 2010.
36.• Imel ZE, Barco JS, Brown HJ, Baucom BR, Baer JS, Kircher JC, et al. The association of therapist empathy and synchrony in vocally encoded arousal. J Couns Psychol. 2014;61(1):146. **This work**

showed the correlation between empathy ratings and the synchrony of therapist-client in fundamental frequency of speech.

37.• Xiao B, Georgiou PG, Imel ZE, Atkins DC, Narayanan SS. Modeling therapist empathy and vocal entrainment in drug addiction counseling. Proc. Interspeech; 2013. **This work showed computational vocal similarity measures that correlate with empathy ratings**.

38. Bishop CM. Pattern recognition and machine learning. Springer; 2007.

39. Kullback S, Leibler RA. On information and sufficiency. Ann Math Stat. 1951;22(1):79–86.

40.•• Xiao B, Imel ZE, Atkins D, Georgiou P, Narayanan SS. Analyzing speech rate entrainment and its relation to therapist empathy in drug addiction counseling. Dresden: Proc. Interspeech; 2015. **This work demonstrated entrainment of speech rate among interlocutors, and found that averaged absolute difference of speech rates between the therapist and client correlates to therapist empathy ratings.**

41. Valstar, M., Girard, J., Almaev, T., McKeown, G., Mehu, M., Yin, L., et al.: FERA 2015-second facial expression recognition and analysis challenge. Proc. IEEE ICFG (2015)

42. Kumano, S., Otsuka, K., Mikami, D., Yamato, J.: Analyzing empathetic interactions based on the probabilistic modeling of the co-occurrence patterns of facial expressions in group meetings. In: Automatic Face and Gesture Recognition, pp. 43–50. IEEE (2011)

43. Murphy KP. Dynamic Bayesian networks: representation, inference and learning, Ph.D. thesis. Berkeley: University of California; 2002.

44.• Kumano S, Otsuka K, Matsuda M, Yamato J. Analyzing perceived empathy based on reaction time in behavioral mimicry. IEICE Trans Inf Syst. 2014;97(8):2008–20. **This work investigated the function of reaction timing in relation to empathy perception, in addition to facial expression and gaze cues.**

45.• Kumano S, Otsuka K, Mikami D, Matsuda M, Yamato J. Analyzing interpersonal empathy via collective impressions. IEEE Trans Affect Comput. 2015;99. **This work analyzed empathy based on an ensemble of evaluators and the distribution of perceived empathy levels**.

46. Tapus, A., Mataric, M.J.: Emulating empathy in socially assistive robotics. In: AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics, pp. 93–96 (2007)

47. Riek LD, Robinson P. Real-time empathy: facial mimicry on a robot. In: Workshop on affective interaction in natural environments (AFFINE) at the international ACM conference on multimodal interfaces (ICMI 08). Citeseer: ACM; 2008.

48. Gonsior, B., Sosnowski, S., Mayer, C., Blume, J., Radig, B., Wollherr, D., et al.: Improving aspects of empathy and subjective performance for HRI through mirroring facial expressions. In: RO-MAN, pp. 350–356. IEEE (2011)

49. Becker C, Prendinger H, Ishizuka M, Wachsmuth I. Evaluating affective feedback of the 3d agent max in a competitive cards game. In: Affective computing and intelligent interaction. Springer; 2005. p. 466–73.

50. Boukricha, H., Wachsmuth, I.: Mechanism, modulation, and expression of empathy in a virtual human. In: IEEE Workshop on Affective Computational Intelligence (WACI), pp. 1–8 (2011)

51.• Boukricha H, Wachsmuth I, Carminati MN, Knoeferle P. A computational model of empathy: empirical evaluation. In: Proc. ACII. IEEE; 2013. p. 1–6. **This work proposed a 3-dim emotion representation space and a three step empathy simulation approach, including empathy mechanism, modulation, and expression**.

52.• Ochs M, Sadek D, Pelachaud C. A formal model of emotions for an empathic rational dialog agent. Auton Agent Multi-Agent Syst. 2012;24(3):410–40. **This work proposed an approach of modeling empathy simulation through formal language and logic on emotion and intention of the user and the virtual agent.**

53. Rodrigues S, Mascarenhas S, Dias J, Paiva A. "I can feel it too!": emergent empathic reactions between synthetic characters. Proc. ACII; 2009. p. 1–7.

54.•• Rodrigues SH, Mascarenhas S, Dias JA, Paiva A. A process model of empathy for virtual agents. Interacting with Computers. 2014. **This work improves previous work by composing emotion modeling, elicitation, modulation and expression**.

55. McQuiggan SW, Lester JC. Learning empathy: a data-driven framework for modeling empathetic companion agents. In: Proceedings of the fifth international joint conference on autonomous agents and multiagent systems. New York: ACM; 2006. p. 961–8.

56. McQuiggan SW, Lester JC. Modeling and evaluating empathy in embodied companion agents. Int J Hum Comput Stud. 2007;65(4): 348–60.

57.• Leite I, Pereira A, Castellano G, Mascarenhas S, Martinho C, Paiva A. Modelling empathy in social robotic companions. In: Ardissono L, Kuflik T, editors. Advances in user modeling, *lecture notes in computer science*, vol. 7138. Heidelberg: Springer Berlin; 2012. p. 135–47. **This work simulated empathic behavior on the robot in a chess game by monitoring the user's emotional and contextual states.**

58. Auer P, Cesa-Bianchi N, Fischer P. Finite-time analysis of the multiarmed bandit problem. Mach Learn. 2002;47(2–3):235–56.

59.• Leite I, Pereira A, Mascarenhas S, Martinho C, Prada R, Paiva A. The influence of empathy in human-robot relations. Int J Hum Comput Stud. 2013;71(3):250–60. **This work compared human perception toward an empathy simulating robot and a neutral behaving one, where the former monitors user's contextual states in a chess game so as to react in a way that may be perceived as empathic.**

60.• D'mello S, Graesser A. AutoTutor and affective autotutor: learning by talking with cognitively and emotionally intelligent computers that talk back. ACM Trans Interact Intell Syst. 2013;2(4):23:1–23:39. **This work introduced a virtual agent that monitors the user's behavioral states and react with empathic expressions in a teaching environment.**

61. Barrett-Lennard GT. The empathy cycle: refinement of a nuclear concept. J Couns Psychol. 1981;28(2):91.

62. Baer JS, Wells EA, Rosengren DB, Hartzler B, Beadnell B, Dunn C. Agency context and tailored training in technology transfer: a pilot evaluation of motivational interviewing training for community counselors. J Subst Abus Treat. 2009;37(2):191.

63. Lee CM, Kilmer JR, Neighbors C, Atkins DC, Zheng C, Walker DD, et al. Indicated prevention for college student marijuana use: a randomized controlled trial. J Consult Clin Psychol. 2013;81(4): 702.

64. Lee CM, Neighbors C, Lewis MA, Kaysen D, Mittmann A, Geisner IM, et al. Randomized controlled trial of a spring break intervention to reduce high-risk drinking. J Consult Clin Psychol. 2014;82(2): 189.

65. Neighbors C, Lee CM, Atkins DC, Lewis MA, Kaysen D, Mittmann A, et al. A randomized controlled trial of event-specific prevention strategies for reducing problematic drinking associated with 21st birthday celebrations. J Consult Clin Psychol. 2012;80(5):850.

66. Roy-Byrne P, Bumgardner K, Krupski A, Dunn C, Ries R, Donovan D, et al. Brief intervention for problem drug use in safety-net primary care settings: a randomized clinical trial. JAMA. 2014;312(5):492–501.

67. Tollison SJ, Lee CM, Neighbors C, Neil TA, Olson ND, Larimer ME. Questions and reflections: the use of motivational interviewing microskills in a peer-led brief alcohol intervention for college students. Behav Ther. 2008;39(2):183–94.

68. Suykens JA, Vandewalle J. Least squares support vector machine classifiers. Neural Process Lett. 1999;9(3):293–300.
69. Bone D, Goodwin MS, Black MP, Lee CC, Audhkhasi K, Narayanan S. Applying machine learning to facilitate autism diagnostics: pitfalls and promises. J Autism Dev Disord. 2015;45(5): 1121–36.
70. Can D, Gibson J, Vaz C, Georgiou P, Narayanan S. Barista: a framework for concurrent speech processing by USC-SAIL. Florence, Italy: Proc. ICASSP; 2014.
71. Van Segbroeck M, Tsiartas A, Narayanan SS. A robust frontend for VAD: exploiting contextual, discriminative and spectral cues of human voice. Lyon, France: Proc. InterSpeech; 2013.
72. Huang CW, Xiao B, Georgiou P, Narayanan S. Unsupervised speaker diarization using Riemannian manifold clustering. Proc. Interspeech; 2014.
73. Miro XA, Bozonnet S, Evans N, Fredouille C, Friedland G, Vinyals O. Speaker diarization: a review of recent research. IEEE Trans Audio Speech Lang Process. 2012;20(2):356–70.
74. Hinton G, Deng L, Yu D, Dahl GE, Mohamed AR, Jaitly N, et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. IEEE Signal Process Mag. 2012;29(6):82–97.
75. Povey D, Ghoshal A, Boulianne G, Burget L, Glembek O, Goel N, et al. The Kaldi speech recognition toolkit. Proc. ASRU; 2011.
76. Cole R, Halpern A, Ramig L, Van Vuuren S, Ngampatipatpong N, Yan J. A virtual speech therapist for individuals with Parkinson's disease. Educ Technol. 2007;47(1):51.
77. Kleinsmith A, Rivera-Gutierrez D, Finney G, Cendan J, Lok B. Understanding empathy training with virtual patients. Comput Hum Behav. 2015;52:151–8.
78. Tettegah S, Taylor K, Whang EW, Meistninkas S, Chamot R. Can virtual reality simulations be used as a research tool to study empathy, problems solving and perspective taking of educators?: theory, method and application. In: ACM SIGGRAPH 2006 educators program. New York: ACM; 2006.
79. van Vuuren S, Cherney LR. A virtual therapist for speech and language therapy. In: Intelligent virtual agents. Springer; 2014. p. 438–48.
80.•• Narayanan S, Georgiou P. Behavioral signal processing: deriving human behavioral informatics from speech and language. Proc IEEE. 2013;101(5):1203–33. **This paper is a survey of the Behavioral Signal Processing work that includes introduction of the concepts, example studies, and future research directions.**