# Asghar, A., Noetel, M., Schellekens, M., Parker, P., Antczak, D., Beauchamp, M., Dicke, T., Diezmann, C., Maeder, A., Ntoumanis, N., Yeung, A., Lonsdale, C. (2021). ASystematic Review of Machine Learning for Assessment and Feedback of Treatment Fidelity. *Psychosocial Intervention.* <https://doi.org/10.5093/pi2021a4>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# **Supplementary file 2.** *Search Strategy, and Included Papers*

## **Search Strategy**

## **Keywords to search in databases**

***Participants***: (teacher\* OR coach\* OR nurse\* OR doctor\* OR physician\* OR "general practitioner\*" OR surgeon\* OR psychiatrist\* OR interviewer\* OR clinician\* OR therap\* OR dentist\* OR physiotherapist\* OR chiropractor\* OR psychotherap\* OR psychologist\* OR counselor\* OR counsellor\* OR "social worker\*" OR "care provider\*")

***Measurement***: (analys\* OR analyz\* OR asses\* OR evaluat\* OR classif\* OR coding OR code OR coded OR rating OR rate\* OR annotat\*)

***Automated coding method***: ("language processing" OR "artificial intelligence" OR "text min\*" OR "data min\*" OR "machine learning" OR "automatic prediction" OR "text classif\*" OR "markov model\*" OR "topic model\*" OR "entropy model\*" OR "language model\*" OR "computational intelligence" OR "recursive neural network\*" OR "discrete sentence feature\*" OR "latent dirichlet allocation" OR "long short term memory")

***Type of behaviour***: (behavio\* OR interaction\* OR interpersonal OR relation\* OR communication\* OR conversation\* OR fidelity OR integrity)

We conducted the main search on 26th Feb 2019.

We updated the search on 21st Feb 2021.

## **Database Search Strategy**

**PubMed Search Strategy**

**PubMed MeSH terms**:

***Participants***: Educational Personnel OR Health Personnel OR Health Occupations

**Measurement**: -

***Automated coding method*:** Natural Language Processing OR Artificial Intelligence

***Type of the behaviour***: Verbal Behavior OR Interpersonal Relations OR Professional-Patient Relations

Main search terms in title and abstract, MeSH terms in “MeSH Major Topic”.

Results: 1,586 records

Search update: 1104 records

**Education Source Search strategy**

**Subject headings**:

***Participants***: Teachers OR coaches OR Health occupations OR Interviewers OR Allied health personnel OR Social workers

***Measurement***: Interaction analysis in education OR Teacher evaluation standards

***Automated coding method*:** Artificial intelligence OR Algorithms

***Type of behaviour***: Verbal behavior OR Interpersonal relations OR Interpersonal communication OR Teacher-student relationships OR Teacher-student communication

Main search terms in title and abstract, Subject Headings in “Subject”.

Results: 198 records

Search update: 81 records

**CINAHL Complete Search Strategy**

**Subject headings**:

***Participants***: Health Personnel OR Social Workers

**Measurement**: -

***Automated coding method***: Natural Language Processing OR Artificial Intelligence

***Type of behaviour***: Interpersonal Relations OR Verbal Behavior

Main search terms in title and abstract, Subject Headings in “Exact Major Subject Heading”.

Results: 297 records

Search update: 271 records

**ERIC Search Strategy**

**Thesaurus**:

***Participants***: Health Occupations OR Health Personnel

***Measurement***: Counselor Evaluation OR Classification OR Teacher Evaluation

**Automated coding method**: Natural Language Processing OR Artificial Intelligence

***Type of behaviour***: Interpersonal Relationship OR Verbal Communication OR Interpersonal Communication

Main search terms in title and abstract, Subject Headings in “Descriptors (exact)”.

Results: 161 records

Search update: 32 records

**PsycINFO Search Strategy**

**Subject headings**:

***Participants***: Educational Personnel OR Health Personnel OR Professional Personnel OR Social Workers

***Measurement***: Behavior Analysis OR Classification

***Automated coding method***: Artificial Intelligence OR Automated Information Coding OR Automated Information Processing

***Type of behaviour***: Verbal Communication OR Teacher Student Interaction OR Interpersonal Interaction OR Interpersonal Relationships OR Interpersonal Communication

Main search terms in title and abstract, Subject Headings in “MeSH Subject Headings”.

Results: 373 records

Search update: 148 Records

**SPORTDiscus search strategy**

**Subject headings:**

***Participants***: TEACHERS OR “COACHES (Athletics)” OR MEDICAL personnel OR PHYSICAL therapists OR SPORTS psychologists

***Measurement***: CLASSIFICATION

**Automated coding method**: -

***Type of behaviour***: COMMUNICATION in sports OR PHYSICIAN-patient relations

Main search terms in title and abstract, Subject Headings in “Subjects (Descriptors)”.

Results: 18 records

Search update: 11 records

**Embase Classic+Embase Search strategy**

**Subject headings:**

***Participants***: educational personnel OR social worker OR health care personnel

**Measurement**: -

***Automated coding method***: natural language processing OR artificial intelligence

**Behaviour**: -

Main search terms in title and abstract, Subject Headings in “Subject heading”.

Results: 1,645 records

Search update: 602 records

**Scopus Search Strategy**

**Scopus database does not have subject headings**

Main search terms in title and abstract.

Results: 5076 records

Search update: 2903 records

**Computers & applied sciences Search strategy**

**Computers & applied sciences Search strategy database does not have subject headings**

Main search terms in title and abstract.

Results: 193

Search update: 139 records

Total: 9,547

Total records added in the search update: 5,291 records

## **Included Papers**

Althoff, T., Clark, K., & Leskovec, J. (2016). Large-scale Analysis of Counseling Conversations: An Application of Natural Language Processing to Mental Health. *Transactions of the Association for Computational Linguistics*, *4*, 463-476. https://doi.org/10.1162/tacl\_a\_00111

Atkins, D., Steyvers, M., Imel, Z., & Smyth, P. (2014). Scaling up the evaluation of psychotherapy: evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science: IS*, *9*(1), 49. https://doi.org/10.1186/1748-5908-9-49

Blanchard, N., Donnelly, P., Olney, A., Samei, B., Ward, B., Sun, X., Kelly, S., Nystrand, M., & D’Mello, S. K. (2016a). Identifying teacher questions using automatic speech recognition in classrooms. *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 191-201. <https://www.aclweb.org/anthology/W16-3623.pdf>

Blanchard, N., Donnelly, P. J., Olney, A. M., Samei, B., Ward, B., Sun, X., Kelly, S., Nystrand, M., & D’Mello, S. K. (2016b). Semi-automatic detection of teacher questions from human-transcripts of audio in live classrooms. *Proceedings of the 9th International Conference on Educational Data Mining*. https://eric.ed.gov/?id=ED592742

Can, D., Atkins, D. C., & Narayanan, S. S. (2015). *A dialog act tagging approach to behavioral coding: A case study of addiction counseling conversations*. Sixteenth Annual Conference of the International Speech Communication Association. https://www.isca-speech.org/archive/interspeech\_2015/i15\_0339.html

Can, D., Georgiou, P. G., & Atkins, D. C. (2012). A case study: Detecting counselor reflections in psychotherapy for addictions using linguistic features. 13th Annual Conference of International Speech Communication Association. <https://www.isca-speech.org/archive/interspeech_2012/i12_2254.html>

Can, D., Georgiou, P., Atkins, D., Marín, R., Imel, Z., & Narayanan, S. (2016). “It sounds like ...”: A natural language processing approach to detecting counselor reflections in motivational interviewing. *Journal of Counseling Psychology*, *63*(3), 343-350. https://doi.org/10.1037/cou0000111

Cao, J., Tanana, M., Imel, Z., Poitras, E., Atkins, D., & Srikumar, V. (2019). Observing dialogue in therapy: Categorizing and forecasting behavioral codes. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 5599–5611. https://doi.org/10.18653/v1/P19-1563

Chakravarthula, S. N., Xiao, B., Imel, Z. E., Atkins, D. C., & Georgiou, P. G. (2015). Assessing empathy using static and dynamic behavior models based on therapist’s language in addiction counseling. *Sixteenth Annual Conference of the International Speech Communication Association*. https://doi.org/10.7287/peerj-cs.59v0.1/reviews/3

Chen, Z., Singla, K., Gibson, J., Can, D., Imel, Z. E., Atkins, D. C., Georgiou, P., & Narayanan, S. (2019). Improving the prediction of therapist behaviors in addiction counseling by exploiting class confusions. *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6605–6609. https://doi.org/10.1109/ICASSP.2019.8682885

Donnelly, P. J., Blanchard, N., Olney, A. M., Kelly, S., Nystrand, M., & D’Mello, S. K. (2017). Words matter: Automatic detection of teacher questions in live classroom discourse using linguistics, acoustics, and context. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 218-227. https://doi.org/10.1145/3027385.3027417

Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., Kelly, S., Nystran, M., & D’Mello, S. K. (2016a). Automatic teacher modeling from live classroom audio. *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, 45-53. <https://doi.org/10.1145/2930238.2930250>

Donnelly, P. J., Blanchard, N., Samei, B., Olney, A. M., Sun, X., Ward, B., Kelly, S., Nystrand, M., & D’Mello, S. K. (2016b). Multi-sensor modeling of teacher instructional segments in live classrooms. *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, 177-184. https://doi.org/10.1145/2993148.2993158

Flemotomos, N., Martinez, V. R., Gibson, J., Atkins, D. C., Creed, T., & Narayanan, S. S. (2018). Language features for automated evaluation of cognitive behavior psychotherapy sessions. *INTERSPEECH*, 1908-1912. 10.21437/Interspeech.2018-1518

Gallo, C., Pantin, H., Villamar, J., Prado, G., Tapia, M., Ogihara, M., Cruden, G., & Brown, C. H. (2015). Blending qualitative and computational linguistics methods for fidelity assessment: Experience with the Familias Unidas Preventive Intervention. *Administration and Policy in Mental Health*, *42*(5), 574-585. https://doi.org/10.1007/s10488-014-0538-4

Gaut, G., Steyvers, M., Imel, Z., Atkins, D., & Smyth, P. (2017). Content coding of psychotherapy transcripts using labeled topic models. *IEEE Journal of Biomedical and Health Informatics*, *21*(2), 476-487. https://doi.org/10.1109/JBHI.2015.2503985

Gibson, J., Atkins, D., Creed, T., Imel, Z., Georgiou, P., & Narayanan, S. (2019). Multi-label multi-task deep learning for behavioral coding. *IEEE Transactions on Affective Computing*, 1-1. https://doi.org/10.1109/TAFFC.2019.2952113

Gibson, J., Can, D., Georgiou, P., Atkins, D. C., & Narayanan, S. S. (2017). Attention networks for modeling behaviors in addiction counseling. *Interspeech 2017*, 3251-3255. https://doi.org/10.21437/Interspeech.2017-218

Gibson, J., Can, D., Xiao, B., Imel, Z. E., Atkins, D. C., Georgiou, P., & Narayanan, S. S. (2016). A deep learning approach to modeling empathy in addiction counseling. *Interspeech 2016*, *2016*, 1447-1451. https://doi.org/10.21437/Interspeech.2016-554

Goldberg, S. B., Flemotomos, N., & Martinez, V. R. (2020). Machine learning and natural language processing in psychotherapy research: Alliance as example use case. *Journal of Counseling*. https://psycnet.apa.org/record/2020-46941-003

Gupta, R., Georgiou, P. G., Atkins, D. C., & Narayanan, S. S. (2014). Predicting client’s inclination towards target behavior change in motivational interviewing and investigating the role of laughter. *Fifteenth Annual Conference of the International Speech Communication Association*. https://www.isca-speech.org/archive/interspeech\_2014/i14\_0208.html

Hasan, M., Carcone, A. I., Naar, S., Eggly, S., Alexander, G. L., Hartlieb, K. E. B., & Kotov, A. (2019). Identifying effective motivational interviewing communication sequences using automated pattern analysis. *Journal of Healthcare Informatics Research*, *3*(1), 86-106. https://doi.org/10.1007/s41666-018-0037-6

Hasan, M., Kotov, A., Carcone, A. I., Dong, M., & Naar, S. (2018). Predicting the outcome of patient-provider communication sequences using recurrent neural networks and probabilistic models. *AMIA Joint Summits on Translational Science Proceedings. AMIA Joint Summits on Translational Science*, *2017*, 64-73. <https://www.ncbi.nlm.nih.gov/pubmed/29888043>

Hasan, M., Kotov, A., Carcone, A., Dong, M., Naar, S., & Hartlieb, K. B. (2016). A study of the effectiveness of machine learning methods for classification of clinical interview fragments into a large number of categories. *Journal of Biomedical Informatics*, *62*, 21-31. https://doi.org/10.1016/j.jbi.2016.05.004

Howes, C., Purver, M., & McCabe, R. (2013). Using conversation topics for predicting therapy outcomes in schizophrenia. *Biomedical Informatics Insights*, *6*(Suppl 1), 39-50. https://doi.org/10.4137/BII.S11661

Idalski Carcone, A., Hasan, M., Alexander, G. L., Dong, M., Eggly, S., Brogan Hartlieb, K., Naar, S., MacDonell, K., & Kotov, A. (2019). Developing machine learning models for behavioral coding. *Journal of Pediatric Psychology*, *44*(3), 289-299. https://doi.org/10.1093/jpepsy/jsy113

Imel, Z., Steyvers, M., & Atkins, D. (2015). Computational psychotherapy research: Scaling up the evaluation of patient-provider interactions. *Psychotherapy* , *52*(1), 19-30. https://doi.org/10.1037/a0036841

Lacson, R., & Barzilay, R. (2005). Automatic processing of spoken dialogue in the home hemodialysis domain. *AMIA Annual Symposium Proceedings,* 420-424. https://www.scopus.com/inward/record.uri?eid=2-s2.0-39049191868&partnerID=40&md5=c99769d457f6ce7d55d48949fa865b04

Malandrakis, N., & Narayanan, S. S. (2015). Therapy language analysis using automatically generated psycholinguistic norms. *Sixteenth Annual Conference of the International Speech Communication Association*. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84959123480&partnerID=40&md5=48c4f18021e9a3d03bf0493b5f6ee649

Mayfield, E., Laws, M. B., Wilson, I. B., & Penstein Rosé, C. (2014). Automating annotation of information-giving for analysis of clinical conversation. *Journal of the American Medical Informatics Association: JAMIA*, *21*(e1), e122-e128. https://doi.org/10.1136/amiajnl-2013-001898

Mieskes, M., & Stiegelmayr, A. (2019). Preparing data from psychotherapy for natural language processing. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*, 2896-2902). https://www.scopus.com/inward/record.uri?eid=2-s2.0-85059902190&partnerID=40&md5=f539dde0dce50ce71931c25747add595

Nitti, M., Ciavolino, E., Salvatore, S., & Gennaro, A. (2010). Analyzing psychotherapy process as intersubjective sensemaking: An approach based on discourse analysis and neural networks. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *20*(5), 546-563. https://doi.org/10.1080/10503301003641886

Park, J., Jindal, A., Kuo, P., Tanana, M., Lafata, J. E., Tai-Seale, M., Atkins, D. C., Imel, Z. E., & Smyth, P. (2021). Automated rating of patient and physician emotion in primary care visits. *Patient Education and Counseling*. https://doi.org/10.1016/j.pec.2021.01.004

Park, J., Kotzias, D., Kuo, P., Logan, R. L., Iv, Merced, K., Singh, S., Tanana, M., Karra Taniskidou, E., Lafata, J. E., Atkins, D. C., Tai-Seale, M., Imel, Z. E., & Smyth, P. (2019). Detecting conversation topics in primary care office visits from transcripts of patient-provider interactions. *Journal of the American Medical Informatics Association: JAMIA*, *26*(12), 1493-1504. https://doi.org/10.1093/jamia/ocz140

Pérez-Rosas, V., Mihalcea, R., Resnicow, K., Singh, S., An, L., Goggin, K. J., & Catley, D. (2017). Predicting counselor behaviors in motivational interviewing encounters. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, 1128-1137. https://www.aclweb.org/anthology/E17-1106

Perez-Rosas, V., Sun, X., Li, C., Wang, Y., Resnicow, K., & Mihalcea, R. (n.d.). *Analyzing the quality of counseling conversations: The tell-tale signs of high-quality counseling*.

Salvatore, S., Gennaro, A., Auletta, A. F., Tonti, M., & Nitti, M. (2012). Automated method of content analysis: A device for psychotherapy process research. *Psychotherapy Research: Journal of the Society for Psychotherapy Research*, *22*(3), 256-273. https://doi.org/10.1080/10503307.2011.647930

Samei, B., Olney, A. M., Kelly, S., Nystrand, M., & D’Mello, S. (2014). Domain independent assessment of dialogic properties of classroom discourse. *Grantee Submission*. https://eric.ed.gov/?id=ED566380

Samei, B., Olney, A. M., Kelly, S., Nystrand, M., D’Mello, S., Blanchard, N., & Graesser, A. (2015). Modeling classroom discourse: Do models that predict dialogic instruction properties generalize across populations? *International Educational Data Mining Society*. https://eric.ed.gov/?id=ED560879

Sen, T., Ali, M. R., Hoque, M. E., Epstein, R., & Duberstein, P. (2017). Modeling doctor-patient communication with affective text analysis. *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. https://doi.org/10.1109/acii.2017.8273596

Singla, K., Chen, Z., Flemotomos, N., Gibson, J., Can, D., Atkins, D., & Narayanan, S. (2018). Using prosodic and lexical information for learning utterance-level behaviors in psychotherapy. *Interspeech 2018*, 3413-3417. https://doi.org/10.21437/Interspeech.2018-2551

Song, Y., Lei, S., Hao, T., Lan, Z., & Ding, Y. (2020). Automatic classification of semantic content of classroom dialogue. *Journal of Educational Computing Research*, 0735633120968554. https://doi.org/10.1177/0735633120968554

Suresh, A., Sumner, T., Jacobs, J., Foland, B., & Ward, W. (2019). Automating analysis and feedback to improve mathematics teachers’ classroom discourse. *Proceedings of the AAAI Conference on Artificial Intelligence, 33*, 9721-9728. https://doi.org/10.1609/aaai.v33i01.33019721

Tanana, M., Hallgren, K. A., Imel, Z. E., Atkins, D. C., & Srikumar, V. (2016). A Comparison of natural language processing methods for automated coding of motivational interviewing. *Journal of Substance Abuse Treatment*, *65*, 43-50. https://doi.org/10.1016/j.jsat.2016.01.006

Velasquez, P. A. E., & Montiel, C. J. (2018). Reapproaching Rogers: A discursive examination of client-centered therapy. *Person-Centered and Experiential Psychotherapies*, *17*(3), 253-269. https://doi.org/10.1080/14779757.2018.1527243

Wallace, B. C., Laws, M. B., Small, K., Wilson, I. B., & Trikalinos, T. A. (2014). Automatically annotating topics in transcripts of patient-provider interactions via machine learning. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, *34*(4), 503-512. https://doi.org/10.1177/0272989X13514777

Wallace, B. C., Trikalinos, T. A., Laws, M. B., Wilson, I. B., & Charniak, E. (2013). A generative joint, additive, sequential model of topics and speech acts in patient-doctor communication. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 1765-1775. https://www.aclweb.org/anthology/D13-1182

Wang, Z., Pan, X., Miller, K. F., & Cortina, K. S. (2014). Automatic classification of activities in classroom discourse. *Computers & Education*, *78*, 115-123. https://doi.org/10.1016/j.compedu.2014.05.010

Xiao, B., Can, D., Georgiou, P. G., Atkins, D., & Narayanan, S. S. (2012). Analyzing the language of therapist empathy in motivational interview based psychotherapy. *Signal and Information Processing Association Annual Summit and Conference (APSIPA), Asia-Pacific*, *2012*. https://www.ncbi.nlm.nih.gov/pubmed/27602411

Xiao, B., Can, D., Gibson, J., Imel, Z., Atkins, D., Georgiou, P., & Narayanan, S. (2016). Behavioral coding of therapist language in addiction counseling using recurrent neural networks. *Interspeech 2016*, 908-912. https://doi.org/10.21437/Interspeech.2016-1560

Xiao, B., Huang, C., Imel, Z., Atkins, D., Georgiou, P., & Narayanan, S. (2016). A technology prototype system for rating therapist empathy from audio recordings in addiction counseling. *PeerJ. Computer Science*, *2*(4). https://doi.org/10.7717/peerj-cs.59

Xiao, B., Imel, Z. E., Georgiou, P., Atkins, D., & Narayanan, S. (2015). “Rate my therapist”: Automated detection of empathy in drug and alcohol counseling via speech and language processing. *PloS One*, *10*(12), e0143055. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0143055