

Using decision trees to characterize verbal communication during change and stuck episodes in the therapeutic process

Víctor H. Masías^{1,2*}, Mariane Krause³, Nelson Valdés³, J. C. Pérez⁴ and Sigifredo Laengle¹

¹ Department of Management Control and Information Systems, Universidad de Chile, Santiago, Chile, ² Faculty of Economics and Business, Universidad Diego Portales, Santiago, Chile, ³ Psychology School, Pontificia Universidad Católica de Chile, Santiago, Chile, ⁴ Faculty of Psychology, Universidad del Desarrollo, Santiago, Chile

OPEN ACCESS

Edited by:

Francesco Pagnini,
Catholic University of Milan, Italy

Reviewed by:

Eleonora Volpato,
Fondazione Don Carlo Gnocchi, Italy
Mauricio A. Valle,
Universidad de Valparaíso, Chile

*Correspondence:

Victor H. Masías,
Department of Management Control
and Information Systems, Universidad
de Chile, Santiago, Chile, Diagonal
Paraguay # 257, Piso 13, Santiago
8330015, Chile;
Faculty of Economics and Business,
Universidad Diego Portales,
Santiago, Chile
vmasias@fen.uchile.cl

Specialty section:

This article was submitted to
Psychology for Clinical Settings, a
section of the journal *Frontiers in
Psychology*

Received: 22 October 2014

Paper pending published:

04 January 2015

Accepted: 17 March 2015

Published: 09 April 2015

Citation:

Masías VH, Krause M, Valdés N,
Pérez JC and Laengle S (2015) Using
decision trees to characterize verbal
communication during change and
stuck episodes in the therapeutic
process. *Front. Psychol.* 6:379.
doi: 10.3389/fpsyg.2015.00379

Methods are needed for creating models to characterize verbal communication between therapists and their patients that are suitable for teaching purposes without losing analytical potential. A technique meeting these twin requirements is proposed that uses decision trees to identify both change and stuck episodes in therapist-patient communication. Three decision tree algorithms (C4.5, NBTree, and REPTree) are applied to the problem of characterizing verbal responses into change and stuck episodes in the therapeutic process. The data for the problem is derived from a corpus of 8 successful individual therapy sessions with 1760 speaking turns in a psychodynamic context. The decision tree model that performed best was generated by the C4.5 algorithm. It delivered 15 rules characterizing the verbal communication in the two types of episodes. Decision trees are a promising technique for analyzing verbal communication during significant therapy events and have much potential for use in teaching practice on changes in therapeutic communication. The development of pedagogical methods using decision trees can support the transmission of academic knowledge to therapeutic practice.

Keywords: decision trees, significant event, coding system, counseling, pilot teaching method

1. Introduction

The gap that has long existed between clinical research and clinical practice in psychotherapy has been widely documented (Barlow, 1981; Elliott, 1983a; Talley et al., 1994; Goldfried and Wolfe, 1996; Monger, 1998; Kazdin, 2001; Jiménez, 2002; Krause, 2011). One of the problems currently face by researchers in therapeutic communication is finding analytic techniques that have pedagogical potential for teaching clinical knowledge not only to therapy educators but also to practicing therapists and the patients themselves. Academic research has found that therapy processes include Change Episodes (CE) as well as Stuck Episodes (SE), both of which are significant events (Elliott, 1983b; Mahrer and Nadler, 1986; Elliott et al., 1985; Gonçalves et al., 2009). Whereas CE generate the transformation of the client's subjective perspective regarding him- or herself, her problems and symptoms, SE can be characterized by the temporary detention of the client's change process (Krause et al., 2006; Fernández et al., 2012). However, no work has been published on the use of alternative analytical techniques for transmitting this knowledge, which has been documented in various qualitative and quantitative

studies (Brehm and Brehm, 1981; Rice and Greenberg, 1984; Etchegoyen, 1987; Bastine et al., 1989; Grafanaki and McLeod, 1999; Safran and Muran, 2000; Arkowitz, 2002; Timulak and Elliott, 2003; Billow, 2006, 2007; Krause et al., 2006; Miron and Brehm, 2006; Ramírez et al., 2006; Gonçalves et al., 2009; Herrera Salinas et al., 2009; Valdés et al., 2012; Fernández et al., 2012).

Existing research on verbal communication shows that uncovering the communication rules in a therapy process is a multi-dimensional analytical problem. On the one hand, to explore the communication between therapist and patient there are numerous Coding Systems (CS) (Friedlander, 1982; Evans et al., 1984; Cobb and Lieberman, 1987; Lieberman and Cobb, 1987; Mahrer et al., 1988; Stiles, 1992; Wiser and Goldfried, 1996; Connolly et al., 1998; Shaikh et al., 2001; Sirigatti, 2004; Trijsburg et al., 2004; Roussos et al., 2006; Del Piccolo et al., 2011; Rimondini, 2011), which are frameworks that contain a set of variables describing the modes of verbal response occurring in a therapeutic process (Valdés et al., 2010b; Froján Parga et al., 2011). And on the other hand, there are a variety of statistical techniques for analyzing these systems (Mazzi, 2011; Gelo et al., 2013; Mörtl and Gelo, 2015). Finally, there exists a range of theories on interpersonal communication in healthcare settings (Street et al., 2009; Bylund et al., 2012; Wouda and van de Wiel, 2013) and an array of techniques for teaching communication (Berkhof et al., 2011; Bylund et al., 2012).

Studies of CE and SE have used logistic regression to analyze the data gathered during these episodes. Though clinical studies have validated this method of approximation (Harre et al., 1988; Steyerberg et al., 2001), reports have shown that even experienced researchers do not always have the training to properly interpret the results of logistic regression analysis (King et al., 2000; Mood, 2010) or the necessary skills to communicate them (Wouda and van de Wiel, 2012). Thus, there is a lack of alternative pedagogical techniques for transmitting knowledge of verbal communication in therapeutic processes to professionals or students who have relatively little background in statistical modeling. This in turn points up the need for techniques that can create models suitable for teaching purposes that are easy to interpret without losing their analytical potential.

This article proposes the use of decision trees (DT), also called classification and regression trees, to analyze and interpret the communication rules that characterize CE and SE. A DT is “a way to represent rules underlying data with hierarchical, sequential structures that recursively partition the data” (Murthy, 1998, p. 345). It is a technique that learns to recognize patterns in data and has performed well in various areas of application (Laengle, 1992; Quinlan, 1993; Rokach, 2007; Wu et al., 2008). Furthermore, DT’s are used as pedagogical support tools to produce easy-to-interpret models generally (Breiman, 2001b; Jormanainen and Sutinen, 2012; Anaya et al., 2013). However, to our knowledge it has not been employed to characterize the verbal communication that takes place during change or SE in therapeutic processes. Some of the advantages that have prompted us to apply DT’s to the study of this phenomenon are the following (Zhao and Zhang, 2008, p. 1956):

- They are easy to understand.
- They are easily converted to a set of production rules.
- They can classify both categorical and numerical data (but the output attribute must be categorical).
- There are no a priori assumptions about the nature of the data.

In our exploration of the performance of DT’s we apply three different DT generation techniques to the problem of classifying episodes as either CE or SE. Our data source is a linguistic corpus of 8 therapies delivered in individual mode that were coded using the Therapeutic Activity Coding System (TACS) (Valdés et al., 2010b) and CE and SE indicators (Krause et al., 2007). An experiment is conducted to test the three DT models, and the best one forms the basis for a proposed pilot teaching method. This method consists of a series of steps to be used by instructors tasked with introducing students to the recognition of change and SE.

The remainder of the paper is organized as follows. Section 2 provides some basic information on change and SE in the therapeutic process; Section 3 details the experimental method employed; Section 4 sets out the results of the experiments, compares the performance of the different DT tested and presents the 6 steps making up our proposed pilot teaching method; and finally, Section 5 presents our conclusions and discusses some practical implications.

2. Change Episodes and Stuck Episodes in the Therapeutic Process

Much research has been done on the processes of therapeutic communication since the days when Freud, practiced his treatment through words based on techniques learned from Breuer (Breuer and Freud, 1895). Over the last 25 years, research into change processes has been directed at *significant events* or episodes during therapy (Elliott, 1983b; Elliott et al., 1985; Mahrer and Nadler, 1986; Gonçalves et al., 2009), focusing on episodes related to change (Rice and Greenberg, 1984; Bastine et al., 1989; Timulak, 2010; Timulak et al., 2010; Marto, 2012) as well as difficulties during these processes (Ramírez et al., 2006; Herrera Salinas et al., 2009). Events linked with change have received several names such as empowerment events (Timulak and Elliott, 2003), innovative moments (Gonçalves et al., 2009), insight (Rice and Greenberg, 1984), helpful events (Grafanaki and McLeod, 1999), and CE (Bastine et al., 1989; Herrera Salinas et al., 2009; Sánchez, 2012). Also, though less frequently, events related to difficulties in the therapeutic process have been conceptualized as ruptures (Safran and Muran, 2000), refusal (Billow, 2006, 2007), reactance (Brehm and Brehm, 1981; Miron and Brehm, 2006), resistance to change (Arkowitz, 2002), impasse (Etchegoyen, 1987), hindering events (Grafanaki and McLeod, 1999), and SE (Ramírez et al., 2006; Herrera Salinas et al., 2009; Sánchez, 2012), to name a few.

A review of the theoretical and empirical literature confirms that CE and SE are the two types of significant episodes existing in

a therapeutic process (see **Figure 1**). As noted earlier, during a CE a transformation of the client's subjective perspective regarding him—or herself, her problems and symptoms, and the association of these with the environment takes place (Krause et al., 2007). This involves the development of new forms of interpretation and representation. An SE, by contrast, can be seen as the opposite of a CE, being characterized by the temporary detention of the client's change process due to the persistence in ways of understanding, behavior and emotions related to his or her problem. In both types of episodes the focus of the observation is primarily on the patient, regardless of the actions or omissions of the therapist and their possible strengths or weaknesses. While in CE new meanings are constructed, SE are characterized by the lack of construction of new modes of interpretation or representation. From a general perspective, CE and SE are two different moments in which therapist and patient configure their alliance and their therapeutic relationship (Safran et al., 1990; Valdés et al., 2010a; McCarthy et al., 2011).

The two episode types are both about 3 min long and are determined through trained clinical observation. The analytical problem thus consists in identifying what verbal communication attributes characterize these episode types but the pedagogical problem is how to teach a therapist in training to detect them. This is especially challenging because identifying these moments requires long therapeutic experience or the expert knowledge of a psychotherapist, neither of which a student therapist will have. In what follows we describe an experimental setup for identifying the elements of CE and SE using DT.

3. Materials and Methods

To build a model capable of classifying CE and SE we designed the experimental setup (depicted in Supplementary Figure 1), which combines qualitative and quantitative data analysis techniques. The method consists of six phases: process recording, data coding, running experiments, calculating performance measures, evaluation of DT via statistical hypothesis testing, and introducing the best of the DT models into teaching and learning. These phases are described in turn below.

3.1. Process Recording

The analysis unit is the “speaking turns” taken by the therapist and the patient participating in the psychodynamic individual-mode therapy processes conducted for the study. After being told of the scope and objectives of our work, both participants signed consent forms permitting the therapy sessions to be recorded, analyzed and quoted from for research purposes as long as anonymity and confidentiality were maintained. The sessions were held in a room with a one-way mirror to facilitate recording in video as well as audio with 8 trained clinical data analysts acting as observers. The dialogues that took place between the therapist and the patient were transcribed verbatim. The research was approved by the Ethics Committee of the Psychology School belonging to the Pontificia Universidad Católica de Chile and by the Ethics Committee of the Chilean National Fund for Research and Technology (FONDECYT).

3.2. Data and Coding

The coded database is an aggregate linguistic corpus of the 8 successful therapy processes. Of the 1760 speaking turns registered during these processes, 1003 belonged to 23 SE and 757 belonged to 22 CE. The codification of the independent variables (communicative actions) and the dependent variable (type of episode) is described below.

3.2.1. Codification of Independent Variables:

The speaking turns were codified in 31 binary categorical variables representing 31 different communicative actions defined by the TACS system (Valdés et al., 2010b) (see Supplementary Table 1). The 8 clinical data analysts observing the sessions determined the presence or absence of these actions. The presence of an action during a speaking turn was codified as a 1 and the absence of an action as a 0.

3.2.2. Codification of Dependent Variables:

To delimit and codify the CE and SE, the speaking turn in which a CE or SE began had first to be identified. To do this the data analysts used a set of 19 generic CE indicators and 11 qualitative thematic SE indicators (Krause et al., 2007) (see Supplementary Tables 2, 3). A binary nominal variable was defined and assigned

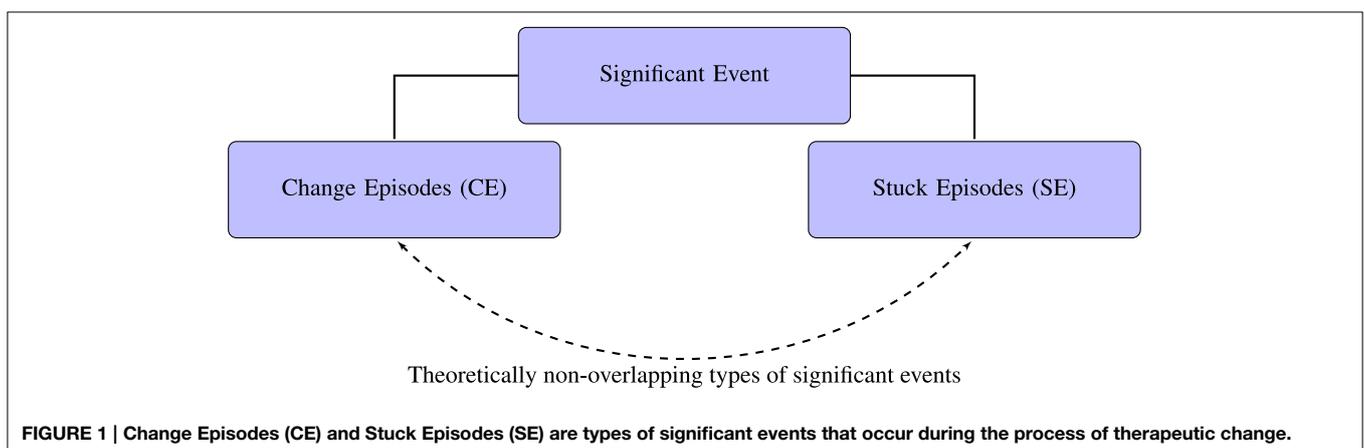


FIGURE 1 | Change Episodes (CE) and Stuck Episodes (SE) are types of significant events that occur during the process of therapeutic change.

a value of 1 for turns occurring during a CE and 0 for turns occurring during an SE.

3.3. Running Experiments

As noted above, three different DT were used to analyze the communicative actions for classifying episodes as either CE or SE. Our first choice was the C4.5 algorithm because the trees it creates are easy to interpret and perform well, but for purposes of comparison we also used the NBTtree and REPTree algorithms (see Supplementary Table 4). To estimate the performance of the DT's, we applied the stratified 10-fold cross-validation approach (Japkowicz and Shah, 2011; Purushotham and Tripathy, 2012), in which "each fold is stratified so that they contain approximately the same proportion of class labels as the original dataset" (Purushotham and Tripathy, 2012, p. 684).

3.4. Performance Measures

The results of the DT models were classified by a confusion matrix (Rokach, 2007) (see Supplementary Figure 2). Based on this matrix, five DT performance measures denoted Overall Accuracy, Precision, Recall, Matthew's Correlation Coefficient (MCC) (Matthews, 1975), and Area Under the Receiver Operating Characteristic curve (ROC Area) were defined by formulas (see Supplementary Table 5). Thus, performance was measured by comparing the values obtained for these indicators.

3.5. Statistical Evaluation of DT's Performance

Two tests were used to evaluate the performance of DT's: the Cochran's Q-Test (Sheskin, 1997) and the McNemar's Test (Bostanci and Bostanci, 2013). The null hypothesis (H_0) of the first test was that they performed similarly whereas the alternative hypothesis (H_1) was that they did not, that is, that they performed differently. If the null hypothesis of Cochran's Q-Test is rejected (i.e., the DT's have different performance), then we applied the second test for each pair of models generated by the algorithms C4.5, NBTtree, and REPTree. Thus, the null hypothesis (H_0) of the McNemar's Test was used to determinate if two DT's have a similar performance, whereas the alternative hypothesis (H_1) was that they did not.

3.6. Teaching and Learning the Best DT Model

Finally, we propose the 6 steps of our pilot teaching method for use by therapy instructors introducing students to the DT model that performed best. These steps were devised by the authors of the present article based on group brainstorming, a qualitative technique for generating ideas that has been used in educational and health professional contexts (Burnard, 1988; Handfield-Jones et al., 1993; Isaksen, 1998; Byron, 2012).

4. Results

This section sets out the performance measure results for the three DT's (Section 4.1), displays the tree generated by the best-performing algorithm and offers presents the steps in the pilot teaching method (both in Section 4.2).

4.1. Comparison and Evaluation of DT's Performance

The performance measure results for the three DT's on the problem of classifying CE and SE speaking turns are summarized in Table 1. As can be seen, the C4.5 algorithm achieved the highest values for the Precision, Recall, ROC Area, and MCC measures. In addition, the DT it generated was smaller and had fewer leaves than the trees produced by the other two algorithms. Note furthermore that although the DT generated by NBTtree performed better than the one created by REPTree, it was also much larger.

To determine whether differences in performance between the algorithms were statistically significant, we conducted the Cochran's Q-Test. The null hypothesis (H_0) was rejected ($Q = 24.43$ with $p = 0.000$), meaning that performances of DT's were statistically different. To determine whether the differences in performance between each pair of algorithms were statistically significant, we conducted the McNemar's Test. The null hypothesis (H_0) was that there were no statistically significant performance differences between two DT's. The null hypothesis was rejected between C4.5 and NBTtree ($T = 16.32$ and $p = 0.000$), and between C4.5 and REPTree ($T = 18.46$ and $p = 0.000$). Additionally, was accepted the null hypothesis between NBTtree and REPTree ($T = 3.23$ and $p = 0.07$). These results show that the performance of C4.5 was indeed statistically different and

TABLE 1 | The Performance of DT's in classifying significant events can be grouped in size of tree (tree size and number of leaves) and prediction capabilities.

Comparison Criteria	Measure	C4.5	NBTtree	REPTree
DT's characteristics	Tree size	29	151	39
	Number of leaves	15	76	20
Performance measures	Correctly classified instances	1,166 66.25%	1,123 63.80%	1,095 62.21%
	Incorrectly classified instances	594 33.75%	637 36.20%	665 37.79%
	Precision	(AVG) 0.71	0.65	0.61
	Recall	(AVG) 0.66	0.63	0.62
	ROC Area	(AVG) 0.66	0.60	0.63
	MCC	(AVG) 0.32	0.24	0.20

superior, while the performance levels of REPTree and NBtree were similar and inferior. Furthermore, the tree generated by C4.5 was smaller than the other two as well as performing better makes it particularly suitable for a pedagogical setting.

4.2. Teaching and Learning Using DT Model

As just suggested above, for teaching purposes it is not only a DT's performance that matters but also its size and the number of rules it requires. This is so because smaller size and fewer leaves mean that the graph a student will have to learn to interpret will have fewer rules and objects. The smaller size and fewer leaves of the C4.5 DT will thus aid in simplifying the interpretability of the model obtained.

The DT graph generated by C4.5 is displayed in **Figure 2**. On the basis of this graph and the group brainstorming process, we drafted the 6 steps in our pilot teaching method for psychotherapy instructors introducing clinical therapy students to this DT. In particular, the brainstorming aimed at generating a method for teaching those with minimal knowledge of statistics to identify EC and SC. The steps themselves are presented in Supplementary Tables 6, 7, 8.

5. Discussion

The purpose of this study was to use the DT technique for analyzing and generating easily interpretable models of verbal

communication during significant moments in therapeutic processes. These processes involve changes in subjective patterns of interpretation and development of new comprehensive theories of oneself (Winkler et al., 1989, 1993; Dagnino et al., 2012), but the construction of such change processes occurs at the level of the conversation between patient and therapist.

Tests were conducted of different DT's before settling on a model that was able to correctly classify 66% of speaking turns as belonging either to a CE or an SE. However, we believe that further research, perhaps using a different CS (see Friedlander, 1982; Evans et al., 1984; Cobb and Lieberman, 1987; Lieberman and Cobb, 1987; Mahrer et al., 1988; Stiles, 1992; Wisner and Goldfried, 1996; Connolly et al., 1998; Shaikh et al., 2001; Sirigatti, 2004; Trijsburg et al., 2004; Roussos et al., 2006; Del Piccolo et al., 2011; Rimondini, 2011), would improve its performance and therefore also that of the model used with the pilot teaching method that was proposed. The idea was to link the development of better classification models to better models of teaching.

One of the interesting findings of our DT model is that Resignifying is the most important variable for classifying a speaking turn as a CE or an SE. In logistic regression analysis there is little consensus and various criteria for evaluating the importance of a predictor (Thomas et al., 2008). By contrast, the DT quickly shows clinical researchers and student therapists which variable is most important. Although it has been previously reported that the Resignifying communicative action is the most frequently

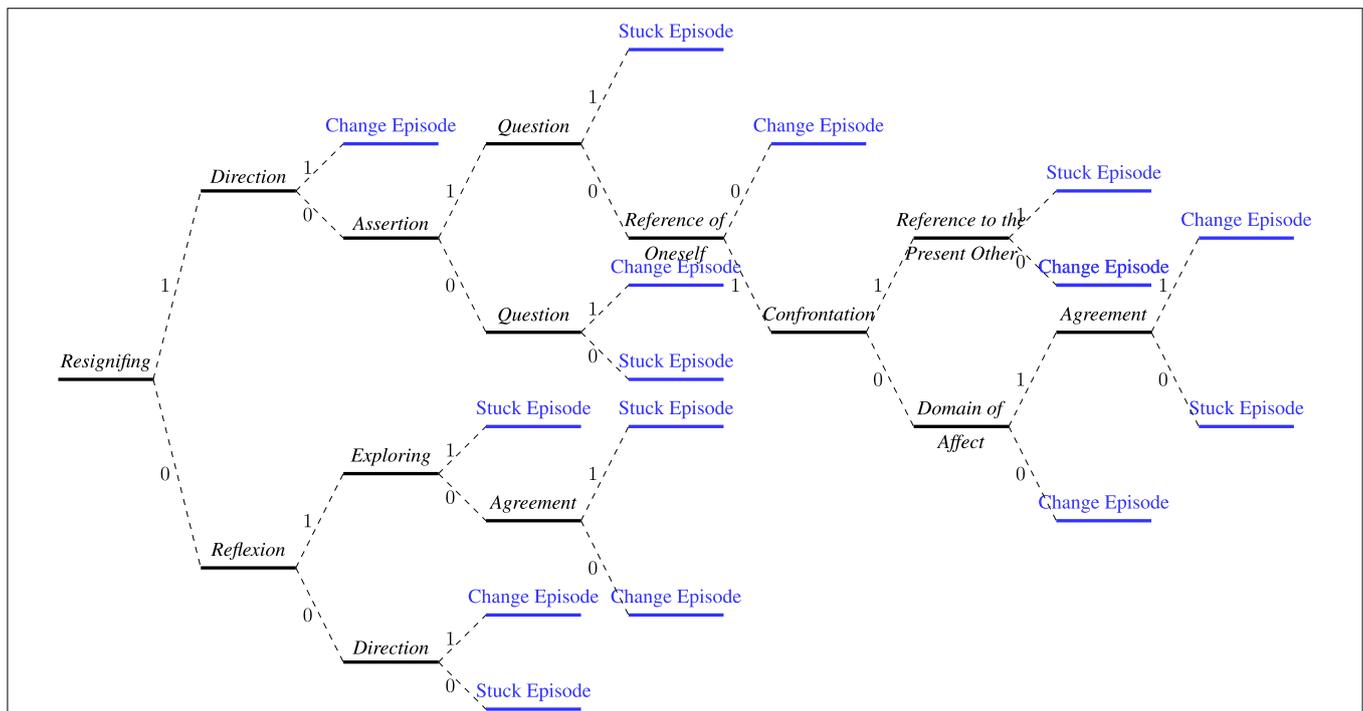


FIGURE 2 | Visualization of decision tree. The black thick line (split nodes) represent communicative actions (independent variables), the dashed lines indicate the variable values (communicative action present equal to 1, communicative action absent equal to 0), and the blue thick line (leaf nodes) indicate the type of episode, that is, CE or SE (dependent variable). Each path from the root node to the leaves is a

communicative rule that classifies speaking turns as CE or SE. Note finally that some formal aspects of the DT have been omitted here in order to focus on the decision rules acquired with the model (exemplified in Supplementary Table 8); further information on the model's statistical properties can be found in Podgorelec et al. (2002), Lee et al. (2009), Perner (2011), and Kotsiantis (2013).

occurring variable in the final phase of therapy (Dagnino et al., 2012), the present study is the first to show that the variable is necessary but not sufficient for classifying CE and SE. Also, the production rules show that Resignifying in the presence of other communicative actions may be characteristic of an SE. This may seem contradictory if we assume that the presence of Resignifying by itself is associated with a CE. Using DT thus shows that the presence of Resignifying in combination with other communicative actions is what distinguishes CE from SE, not Resignifying alone.

The above result indicates how a DT reveals the complexity of the combination of attributes in a therapist-patient conversation. This is consistent with theoretical concepts of psychotherapy according to which the purpose of therapy is to generate new meanings through therapeutic conversation (Watzlawick et al., 1974; Watzlawick, 1976; de Shazer, 1979; Capps, 1990; Pesut, 1991), with resignifying as one of various attributes that contribute to changes. In future research we intend to carry out a pilot study of the application of the proposed teaching method to the measurement of student learning using DT's. Only this way can the potential of DT's for introducing students to the complexity of verbal communication in therapy be empirically tested.

5.1. Practice Implications

Training in communication skills is an educational value (Rotthoff et al., 2011) that is into practice with the available technology. Based on results using data mining, we posit that DT techniques can be introduced into both clinical research on therapeutic communication and the practice of counseling or therapy. By using different scales of measurement as independent variables, DT's are able to analyze data obtained using different existing CS (Evans et al., 1984; Mahrer et al., 1988; Wiser and Goldfried, 1996; Connolly et al., 1998; Shaikh et al., 2001; Sirigatti, 2004; Trijsburg et al., 2004; Roussos et al., 2006), and through its ability to generate a graph of the induced model and production rules, DT's also have potential for use by various actors in addition to researchers who explore issues of communication in therapy. In practical terms, DT techniques are a valuable new pedagogical tool for the study and teaching of verbal communication in therapeutic processes.

5.2. Research Limitations

The method used in this paper has two principal limitations. First, as some studies have demonstrated, DT can be unstable (Last et al., 2002; Kitsantas et al., 2007) with the consequence that labeling and adding new examples to the training set may generate changes in the tree originally obtained. To improve the results presented here a stability analysis of the tree could be undertaken similar to the one conducted in Dwyer and Holte (2007). This would provide information on the tree's stability in addition to the evaluation of its performance already done here.

The second limitation stems from the fact that there exists a variety of measures for analyzing variable importance (Rokach

and Maimon, 2005). For example, the C4.5 algorithm used here applies an approximation heuristic to determine which predictor variable is the most important on the basis of the highest gain ratio. Researchers contemplating extensions of this work could opt for other criteria such as the gain rate or the Gini index (Raileanu and Stoffel, 2004). Another alternative would be to utilize DT algorithms such as Random Forest (Breiman, 2001a) that incorporate a more robust variable importance measure, but this may involve important tradeoffs. In the Random Forest case, the final output is constructed with 500 trees and is thus not humanly readable. In effect, the algorithm is a *black box* in which what is gained in robustness is lost in interpretability.

5.3. Conclusion

Based on the exploratory and comparative results of our study, we conclude that DT techniques have great potential for classifying the modes of verbal responses in therapeutic communication into CE or SE. Greater accuracy may be obtainable through further research into the performance of other CS in classification or other problems. In either case, the techniques considered must take into account the needs of teaching and practical learning. The ultimate goal is to find methods that recognize and reinforce the fundamental concept that therapy using treatment through words is based on a unified paradigm of teaching, learning and research.

Author Contributions

VM designed the research, VM and SL conducted the empirical tests. MK, NV, and JCP carried out the data collection and studied the research domain. VM, MK, NV, and JCP, and SL wrote the paper. All authors read and approved the final manuscript.

Acknowledgments

The authors wish to thank all those who reviewed this paper, and are especially grateful to everyone in the *Programa de Investigación Psicoterapia y Cambio* at the Pontificia Universidad Católica de Chile (PUC) for their enthusiastic support of this project. This study was supported by the Fund for Innovation and Competitiveness (FIC) of the Chilean Ministry of Economy, Development and Tourism, through the Millennium Scientific Initiative, Grant N° IS130005. This work was partially supported by the Faculty of Economics and Business, Universidad Diego Portales, Manuel Rodríguez Sur 253, 8370057, Santiago, Chile.

Supplementary Material

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fpsyg.2015.00379/abstract>

References

- Anaya, A. R., Luque, M., and García-Saiz, T. (2013). Recommender system in collaborative learning environment using an influence diagram. *Expert Syst. Appl.* 40, 7193–7202. doi: 10.1016/j.eswa.2013.07.030
- Arkowitz, H. (2002). Toward an integrative perspective on resistance to change. *J. Clin. Psychol.* 58, 219–227. doi: 10.1002/jclp.1145
- Barlow, D. H. (1981). On the relation of clinical research to clinical practice: current issues, new directions. *J. Consult. Clin. Psychol.* 49, 147.
- Bastine, R., Fiedler, P., and Kommer, D. (1989). Was ist therapeutisch an der Psychotherapie? Versuch einer Vestandsaufnahme und Systematisierung der Psychotherapeutischen Prozeßforschung. *Z. Klin. Psychol.* 18, 3–22.
- Berkhof, M., van Rijssen, H. J., Schellart, A. J., Anema, J. R., and van der Beek, A. J. (2011). Effective training strategies for teaching communication skills to physicians: an overview of systematic reviews. *Patient Educ. Couns.* 84, 152–162. doi: 10.1016/j.pec.2010.06.010
- Billow, R. M. (2006). The three R's of group: resistance, rebellion, and refusal. *Int. J. Group Psychother.* 56, 259–284. doi: 10.1521/ijgp.2006.56.3.259
- Billow, R. M. (2007). On refusal. *Int. J. Group Psychother.* 57, 419–449. doi: 10.1521/ijgp.2007.57.4.419
- Bostanci, B., and Bostanci, E. (2013). “An evaluation of classification algorithms using McNemar's test,” in *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)* (Gwalior: Springer), 15–26.
- Brehm, S. S., and Brehm, J. W. (1981). *Psychological Reactance: a Theory of Freedom and Control*. New York, NY: Academic Press.
- Breiman, L. (2001a). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Breiman, L. (2001b). Statistical modeling: the two cultures (with comments and a rejoinder by the author). *Stat. Sci.* 16, 199–231. doi: 10.1214/ss/1009213726
- Breuer, J., and Freud, S. (1895). *Studien über Hysterie*. Leipzig: F. Deuticke.
- Burnard, P. (1988). Brainstorming: a practical learning activity in nurse education. *Nurse Educ. Today* 8, 354–358.
- Bylund, C. L., Peterson, E. B., and Cameron, K. A. (2012). A practitioner's guide to interpersonal communication theory: an overview and exploration of selected theories. *Patient Educ. Couns.* 87, 261–267. doi: 10.1016/j.pec.2011.10.006
- Byron, K. (2012). Creative reflections on brainstorming. *Lond. Rev. Educ.* 10, 201–213. doi: 10.1080/14748460.2012.691284
- Capps, D. (1990). *Reframing: a New Method in Pastoral Care*. Minneapolis, MN: Fortress Press.
- Cobb, J. P., and Lieberman, S. (1987). The grammar of psychotherapy. a descriptive account. *Br. J. Psychiatry* 151, 589–594.
- Connolly, M. B., Crits-Christoph, P., Shappell, S., Barber, J. P., and Luborsky, L. (1998). Therapist interventions in early sessions of brief supportive-expressive psychotherapy for depression. *J. Psychother. Pract. Res.* 7, 290.
- Dagnino, P., Krause, M., Pérez, C., Valdés, N., and Tomicic, A. (2012). The evolution of communicative intentions during change episodes and throughout the therapeutic process. *Res. Psychother. Psychopathol. Process Outcome* 15, 75–86. doi: 10.7411/RP.2012.008
- de Shazer, S. (1979). On transforming symptoms: an approach to an Erickson procedure. *Am. J. Clin. Hypn.* 22, 17–28.
- Del Piccolo, L., de Haes, H., Heaven, C., Jansen, J., Verheul, W., Bensing, J., et al. (2011). Development of the Verona coding definitions of emotional sequences to code health providers' responses (VR-CoDES-P) to patient cues and concerns. *Patient Educ. Couns.* 82, 149–155. doi: 10.1016/j.pec.2010.02.024
- Dwyer, K., and Holte, R. (2007). “Decision tree instability and active learning,” in *Proceedings of the 18th European Conference on Machine Learning, ECML '07* (Berlin; Heidelberg: Springer-Verlag), 128–139.
- Elliott, R. (1983a). Fitting process research to the practicing psychotherapist. *Psychother. Theory Res. Pract.* 20:47.
- Elliott, R. (1983b). That in your hands. a comprehensive process analysis of a significant event in psychotherapy. *Psychiatry* 46, 113–129.
- Elliott, R., James, E., Reimschuessel, C., Cislo, D., and Sack, N. (1985). Significant events and the analysis of immediate therapeutic impacts. *Psychother. Theory Res. Pract. Train.* 22:620.
- Etchegoyen, R. H. (1987). *Fundamentos da Técnica Psicanalítica*. Buenos Aires: Amorrortu.
- Evans, M. D., Piasecki, J. M., Kriss, M. R., and Hollon, S. D. (1984). *Raters' Manual for the Collaborative Study Psychotherapy Rating Scale, Form 6 (CSPRS-6)*. Minneapolis: University of Minnesota and the St. Paul–Ramsey Medical Center. Available online at: US Department of Commerce, National Technical Information Service, Springfield, VA.
- Fernández, O., Herrera, P., Krause, M., Pérez, J. C., Valdés, N., Vilches, O., et al. (2012). Episodios de cambio y estancamiento en psicoterapia: características de la comunicación verbal entre pacientes y terapeutas. *Terapia Psicol.* 30, 5–22. doi: 10.4067/S0718-48082012000200001
- Friedlander, M. L. (1982). Counseling discourse as a speech event: revision and extension of the Hill counselor verbal response category system. *J. Couns. Psychol.* 29, 425.
- Froján Parga, M. X., Montaña Fidalgo, M., Calero Elvira, A., and Ruiz Sancho, E. (2011). Aproximación al estudio funcional de la interacción verbal entre terapeuta y cliente durante el proceso terapéutico. *Clín. Salud* 22, 69–85.
- Gelo, O. C. G., Salcuni, S., and Colli, A. (2013). Text analysis within quantitative and qualitative psychotherapy process research: introduction to special issue. *Res. Psychother. Psychopathol. Process Outcome* 15, 45–53. doi: 10.7411/RP.2012.005
- Goldfried, M. R., and Wolfe, B. E. (1996). Psychotherapy practice and research: repairing a strained relationship. *Am. Psychol.* 51, 1007.
- Gonçalves, M. M., Matos, M., and Santos, A. (2009). Narrative therapy and the nature of innovative moments in the construction of change. *J. Construct. Psychol.* 22, 1–23. doi: 10.1080/10720530802500748
- Grafanaki, S., and McLeod, J. (1999). Narrative processes in the construction of helpful and hindering events in experiential psychotherapy. *Psychother. Res.* 9, 289–303.
- Handfield-Jones, R., Nasmith, L., Steinert, Y., and Lawn, N. (1993). Creativity in medical education: the use of innovative techniques in clinical teaching. *Med. Teach.* 15, 3–10.
- Harre, F. E., Lee, K. L., and Pollock, B. G. (1988). Regression models in clinical studies: determining relationships between predictors and response. *J. Natl. Cancer Inst.* 80, 1198–1202.
- Herrera Salinas, P., Fernández González, O., Krause Jacob, M., Vilches Álvarez, O., Valdés, N., and Dagnino, P. (2009). Revisión teórica y metodológica de las dificultades en psicoterapia: propuesta de un modelo ordenador. *Terapia Psicol.* 27, 169–179. doi: 10.4067/S0718-48082009000200003
- Isaksen, S. G. (1998). *A Review of Brainstorming Research: Six Critical Issues for Enquiry (Monograph# 302)*. Buffalo, NY: Creative Problem Solving Group–Buffalo.
- Japkowicz, N., and Shah, M. (2011). *Evaluating Learning Algorithms: a Classification Perspective*. Cambridge: Cambridge University Press.
- Jiménez, J. P. (2002). Miradas sobre los puentes entre la práctica y la investigación en psicoterapia. *Terapia Psicol.* 20, 61–74.
- Jormanainen, I., and Sutinen, E. (2012). “Using data mining to support teacher's intervention in a robotics class,” in *Proceedings of the 2012 IEEE Fourth International Conference On Digital Game And Intelligent Toy Enhanced Learning* (Los Alamitos, CA: IEEE Computer Society), 39–46.
- Kazdin, A. E. (2001). Bridging the enormous gaps of theory with therapy research and practice. *J. Clin. Child Psychol.* 30, 59–66. doi: 10.1207/S15374424JCCP3001_7
- King, G., Tomz, M., and Wittenberg, J. (2000). Making the most of statistical analyses: improving interpretation and presentation. *Am. J. Polit. Sci.* 44, 347–361. doi: 10.2307/2669316
- Kitsantas, P., Hollander, M., and Li, L. M. (2007). Assessing the stability of classification trees using Florida birth data. *J. Stat. Plan. Inference* 137, 3917–3929. doi: 10.1016/j.jspi.2007.04.010
- Kotsiantis, S. B. (2013). Decision trees: a recent overview. *Artif. Intell. Rev.* 39, 261–283. doi: 10.1007/s10462-011-9272-4
- Krause, M. (2011). Psychotherapy: practice without science or science without practice? *Rev. Colomb. Psicol.* 20, 89–98.
- Krause, M., De la Parra, G., Aristegui, R., Dagnino, P., Tomicic, A., Valdés, N., et al. (2007). The evolution of therapeutic change studied through generic change indicators. *Psychother. Res.* 17, 673–689. doi: 10.1080/10503300601158814
- Krause, M., Parra, G. D. L., Aristegui, R., Dagnino, P., Tomicic, A., Valdés, N., et al. (2006). Indicadores genéricos de cambio en el proceso psicoterapéutico. *Rev. Latinoam. Psicol.* 38, 299–325.

- Laeagle, S. (1992). *Inducción de Árboles de Decisión*. Master's thesis, Universidad de Chile, Facultad de Ciencias Físicas y Matemáticas. Departamento de Ingeniería Industrial, Santiago, Chile.
- Last, M., Maimon, O., and Minkov, E. (2002). Improving stability of decision trees. *Int. J. Pattern Recogn. Artif. Intell.* 16, 145–159. doi: 10.1142/S0218001402001599
- Lee, A., Joynt, G. M., Ho, A. M., Keitz, S., McGinn, T., and Wyer, P. C. (2009). Tips for teachers of evidence-based medicine: making sense of decision analysis using a decision tree. *J. Gen. Intern. Med.* 24, 642–648. doi: 10.1007/s11606-009-0918-8
- Lieberman, S., and Cobb, J. P. (1987). The grammar of psychotherapy. Interactograms: three self-monitoring instruments for audiotape feedback. *Br. J. Psychiatry* 151, 594–601.
- Mahrer, A. R., and Nadler, W. P. (1986). Good moments in psychotherapy: a preliminary review, a list, and some promising research avenues. *J. Consult. Clin. Psychol.* 54, 10.
- Mahrer, A. R., Nadler, W. P., Stalikas, A., Schachter, H. M., and Sterner, I. (1988). Common and distinctive therapeutic change processes in client-centered, rational-emotive, and experiential psychotherapies. *Psychol. Rep.* 62, 972–974.
- Marto, M. F. F. (2012). *Eventos Significativos no Counselling Psicoespiritual Geradores de Empowerment Para o Processo de Mudança na Vida*. Master's thesis, Lisboa, Portugal: ISPA - Instituto Universitário de Ciências Psicológicas, Sociais e da Vida.
- Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophys. Acta* 405, 442–451.
- Mazzi, M. A. (2011). “Quantitative methods for the analysis of verbal interactions in psychotherapy,” in *Communication in Cognitive Behavioral Therapy*, ed M. Rimondini (New York, NY: Springer), 233–247.
- McCarthy, K. L., Mergenthaler, E., Schneider, S., and Grenyer, B. F. (2011). Psychodynamic change in psychotherapy: cycles of patient-therapist linguistic interactions and interventions. *Psychother. Res.* 21, 722–731. doi: 10.1080/10503307.2011.615070
- Miron, A. M., and Brehm, J. W. (2006). Reactance theory-40 years later. *Z. Sozialpsychol.* 37, 9–18. doi: 10.1024/0044-3514.37.1.9
- Monger, J. (1998). The gap between theory and practice: a consideration of the fee. *Psychodyn. Couns.* 4, 93–106.
- Mood, C. (2010). Logistic regression: why we cannot do what we think we can do, and what we can do about it. *Eur. Sociol. Rev.* 26, 67–82. doi: 10.1093/esr/jcp006
- Mörkl, K., and Gelo, O. C. G. (2015). “Qualitative methods in psychotherapy process research,” in *Psychotherapy Research*, eds O. C. G. Gelo, A. Pritz, and B. Rieken (Vienna: Springer), 381–428.
- Murthy, S. K. (1998). Automatic construction of decision trees from data: a multi-disciplinary survey. *Data Mining Knowl. Discov.* 2, 345–389.
- Perner, P. (2011). “How to interpret decision trees?,” in *Proceedings of the 11th International Conference on Advances in Data Mining: Applications and Theoretical Aspects* (New York, NY: Springer-Verlag), 40–55.
- Pesut, D. J. (1991). The art, science, and techniques of reframing in psychiatric mental health nursing. *Issues Ment. Health Nurs.* 12, 9–18.
- Podgorelec, V., Kokol, P., Stiglic, B., and Rozman, I. (2002). Decision trees: an overview and their use in medicine. *J. Med. Syst.* 26, 445–463. doi: 10.1023/A:1016409317640
- Purushotham, S., and Tripathy, B. (2012). “Evaluation of classifier models using stratified tenfold cross validation techniques,” in *Global Trends in Information Systems and Software Applications* (Vellore: Springer), 680–690.
- Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.
- Raileanu, L. E., and Stoffel, K. (2004). Theoretical comparison between the Gini index and Information Gain criteria. *Ann. Math. Artif. Intell.* 41, 77–93. doi: 10.1023/B:AMAI.0000018580.96245.c6
- Ramirez, I., Krause, M., Reyes, L., Echavarrri, M., Tomicic, A., de la Parra, G., et al. (2006). “No todo es cambio: hacia una definición de episodios de estancamiento en psicoterapia,” in *Ponencia Presentada en el VII Congreso Chileno de Investigación Empírica en Psicoterapia, Capítulo Chileno SPR y Sonapsyn*, (Reñaca, Chile).
- Rice, L. N., and Greenberg, L. S. (1984). *Patterns of Change: Intensive Analysis of Psychotherapy Process*. New York, NY: Guilford Press.
- Rimondini, M. (2011). *Communication in Cognitive Behavioral Therapy*. New York, NY: Springer.
- Rokach, L. (2007). *Data Mining with Decision Trees: Theory and Applications*, Vol. 69. London: World Scientific.
- Rokach, L., and Maimon, O. (2005). Top-down induction of decision trees classifiers—a survey. *IEEE Trans. Syst. Man Cybern. C Appl. Rev.* 35, 476–487. doi: 10.1109/TSMCC.2004.843247
- Rothhoff, T., Baehring, T., David, D. M., Bartnick, C., Linde, F., Willers, R., et al. (2011). The value of training in communication skills for continuing medical education. *Patient Educ. Couns.* 84, 170–175. doi: 10.1016/j.pec.2010.07.034
- Roussos, A. J., Etchebarne, I., and Waizmann, V. (2006). Un esquema clasificatorio para las intervenciones en psicoterapia cognitiva y psicoanalítica. *Anu. Invest.* 13, 51–61. Available online at: http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1851-16862006001010005&lng=es&nrm=iso
- Safra, J. D., Crocker, P., McMain, S., and Murray, P. (1990). Therapeutic alliance rupture as a therapy event for empirical investigation. *Psychother. Theory Res. Pract. Train.* 27, 154.
- Safra, J. D., and Muran, J. C. (2000). Resolving therapeutic alliance ruptures: diversity and integration. *J. Clin. Psychol.* 56, 233–243. doi: 10.1002/(SICI)1097-4679(200002)56:2<233::AID-JCLP9>3.0.CO;2-3
- Sánchez, N. V. (2012). Analysis of verbal emotional expression in change episodes and throughout the psychotherapeutic process: main communicative patterns used to work on emotional contents. *Clin. Salud* 23, 153–179. doi: 10.5093/cl2012a10
- Shaikh, A., Knobloch, L. M., and Stiles, W. B. (2001). The use of a verbal response mode coding system in determining patient and physician roles in medical interviews. *Health Commun.* 13, 49–60. doi: 10.1207/S15327027HC1301_05
- Sheskin, D. J. (1997). *Handbook of Parametric and Nonparametric Statistical Procedures*. New York, NY: CRC Press.
- Sirigatti, S. (2004). Application of the Jones' psychotherapy process Q-sort. *Brief Strateg. Syst. Ther. Eur. Rev.* 1, 194–207.
- Steyerberg, E. W., Harrell, F. E. Jr., Borsboom, G. J., Eijkemans, M., Vergouwe, Y., and Habbema, J. D. F. (2001). Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J. Clin. Epidemiol.* 54, 774–781. doi: 10.1016/S0895-4356(01)00341-9
- Stiles, W. B. (1992). *Describing Talk: a Taxonomy of Verbal Response Modes*. Newbury Park: Sage Publications.
- Street, R. L. Jr., Makoul, G., Arora, N. K., and Epstein, R. M. (2009). How does communication health? pathways linking clinician-patient communication to health outcomes. *Patient Educ. Couns.* 74, 295–301. doi: 10.1016/j.pec.2008.11.015
- Talley, P. F., Strupp, H. H., and Butler, S. F. (1994). *Psychotherapy Research and Practice: Bridging the Gap*. New York, NY: Basic Books.
- Thomas, D. R., Zhu, P., Zumbo, B. D., and Dutta, S. (2008). On measuring the relative importance of explanatory variables in a logistic regression. *J. Mod. Appl. Stat. Methods* 7, 21–38. Available online at: <http://digitalcommons.wayne.edu/jmasm/vol7/iss1/4>
- Timulak, L. (2010). Significant events in psychotherapy: an update of research findings. *Psychol. Psychother. Theory Res. Pract.* 83, 421–447. doi: 10.1348/147608310X499404
- Timulak, L., Belicova, A., and Miler, M. (2010). Client identified significant events in a successful therapy case: the link between the significant events and outcome. *Couns. Psychol. Q.* 23, 371–386. doi: 10.1080/09515070.2010.534329
- Timulak, L., and Elliott, R. (2003). Empowerment events in process-experiential psychotherapy of depression: an exploratory qualitative analysis. *Psychother. Res.* 13, 443–460. doi: 10.1093/ptr/kpg043
- Trijsburg, R. W., Lietaer, G., Colijn, S., Abrahamse, R. M., Joosten, S., and Duivenvoorden, H. J. (2004). Construct validity of the comprehensive psychotherapeutic interventions rating scale. *Psychother. Res.* 14, 346–366. doi: 10.1093/ptr/kph029
- Valdés, N., Dagnino, P., Krause, M., Pérez, J. C., Altimir, C., Tomicic, A., et al. (2010a). Analysis of verbalized emotions in the psychotherapeutic dialogue during change episodes. *Psychother. Res.* 20, 136–150. doi: 10.1080/10503300903170921
- Valdés, N., Krause, M., Tomicic, A., and Espinosa, D. (2012). Expresión emocional verbal durante episodios de cambio: Análisis de los patrones comunicacionales utilizados por pacientes y terapeutas para trabajar contenidos emocionales. *Rev. Argent. Clín. Psicol.* 21, 217–246.

- Valdés, N., Tomicic, A., Pérez, C., and Krause, M. (2010b). Sistema de codificación de la actividad terapéutica (SCAT-1.0): dimensiones y categorías de las acciones comunicacionales de pacientes y psicoterapeutas. *Rev. Argent. Clín. Psicol.* 19, 117–130.
- Watzlawick, P. (1976). “The psychotherapeutic technique of ‘reframing,’” in *Successful Psychotherapy*, ed J. L. Claghorn (New York, NY: Brunner/Mazel), 119–127.
- Watzlawick, P., Weakland, J. H., and Fisch, R. (1974). *Change: Principles of Problem Formation and Problem Resolution*. New York, NY: WW Norton.
- Winkler, M. I., Avendaño, C., Krause, M., and Soto, A. (1993). El cambio psicológico desde la perspectiva de los consultantes. *Terapia Psicol.* 20, 7–19.
- Winkler, M. I., Cáceres, C., Fernández, I., and Sanhueza, J. (1989). Factores inespecíficos de la psicoterapia y efectividad del proceso terapéutico: una sistematización. *Terapia Psicol.* 8, 34–40.
- Wiser, S., and Goldfried, M. (1996). Verbal interventions in significant psychodynamic–interpersonal and cognitive–behavioral therapy sessions. *Psychother. Res.* 6, 309–319.
- Wouda, J. C., and van de Wiel, H. (2012). The communication competency of medical students, residents and consultants. *Patient Educ. Couns.* 86, 57–62. doi: 10.1016/j.pec.2011.03.011
- Wouda, J. C., and van de Wiel, H. (2013). Education in patient–physician communication: how to improve effectiveness? *Patient Educ. Couns.* 90, 46–53. doi: 10.1016/j.pec.2012.09.005
- Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., et al. (2008). Top 10 algorithms in data mining. *Knowl. Inf. Syst.* 14, 1–37. doi: 10.1007/s10115-007-0114-2
- Zhao, Y., and Zhang, Y. (2008). Comparison of decision tree methods for finding active objects. *Adv. Space Res.* 41, 1955–1959. doi: 10.1016/j.asr.2007.07.020
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Masías, Krause, Valdés, Pérez and Laengle. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.