Social Science & Medicine 198 (2018) 22–26

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Short communication

Latent Semantic Analysis: A new measure of patient-physician communication

Scott R. Vrana[a,*], Dylan T. Vrana[b], Louis A. Penner[c], Susan Eggly[d], Richard B. Slatcher[e], Nao Hagiwara[a]

[a] Department of Psychology, 806 W. Franklin St. Box 84-2018, Virginia Commonwealth University, Richmond, VA 23284, USA
[b] Ray and Stephanie Lane Center for Computational Biology, Carnegie Mellon University, 5000 Forbes Ave Suite 7725, Pittsburgh, PA 15213, USA
[c] Wayne State University, Population Studies and Disparities Research Program, Karmanos Cancer Institute, Detroit, MI 48201, USA
[d] Department of Oncology, Karmanos Cancer Institute/Wayne State University, 4100 John R, MM03CB, Detroit, MI 48201, USA
[e] Department of Psychology, Wayne State University, 5057 Woodward Ave., Detroit, MI 48202, USA

ARTICLE INFO

Keywords:
Latent Semantic Analysis
Patient-physician communication
Communication
Text analysis
Methodology

ABSTRACT

Rationale: Patient–physician communication plays an essential role in a variety of patient outcomes; however, it is often difficult to operationalize positive patient-physician communication objectively, and the existing evaluation tools are generally time-consuming.
Objective: This study proposes semantic similarity of the patient's and physician's language in a medical interaction as a measure of patient-physician communication. Latent semantic analysis (LSA), a mathematical method for modeling semantic meaning, was employed to assess similarity in language during clinical interactions between physicians and patients.
Methods: Participants were 132 Black/African American patients (76% women, *M*age = 43.8, range = 18–82) who participated in clinical interactions with 17 physicians (53% women, *M*age = 27.1, range = 26–35) in a primary care clinic in a large city in the Midwestern United States.
Results: LSA captured reliable information about patient-physician communication: The mean correlation indicating similarity between the transcripts of a physician and patient in a clinical interaction was 0.142, significantly greater than zero; the mean correlation between a patient's transcript and transcripts of their physician during interactions with other patients was not different from zero. Physicians differed significantly in the semantic similarity between their language and that of their patients, and these differences were related to physician ethnicity and gender. Female patients exhibited greater communication similarity with their physicians than did male patients. Finally, greater communication similarity was predicted by less patient trust in physicians prior to the interaction and greater patient trust after the interaction.
Conclusion: LSA is a potentially important tool in patient-physician communication research. Methodological considerations in applying LSA to address research questions in patient-physician communication are discussed.

1. Introduction

Patient–physician communication plays an essential role in a variety of patient outcomes, ranging from trust in health care to treatment adherence, and ultimately to health outcomes (Epstein and Street, 2007; Matusitz and Spear, 2014; Ong et al., 1995; Stewart, 1995). The definition of high-quality patient-physician communication differs between two major frameworks: patient-centered communication and relationship-centered communication, with the former placing more focus on the role of physicians in listening to, informing, and involving patients in their care (Institute of Medicine, 2001) whereas the latter places more focus on reciprocal influences between physicians and patients (Roter, 2000). However, both frameworks emphasize the importance of responsiveness between physicians and patients (Beach and Inui, 2006; Davis et al., 2005; Epstein and Street, 2011; Stewart et al., 2003; Suchman, 2006).

There are multiple ways to operationalize responsiveness between a physician and a patient (see Boon and Stewart, 1998; Epstein et al., 2005 for review), including patient and doctor questionnaires about the interaction, observational techniques, and analyses of transcribed

* Corresponding author.
E-mail addresses: srvrana@vcu.edu (S.R. Vrana), dtv@andrew.cmu.edu (D.T. Vrana), pennerl@karmanos.org (L.A. Penner), egglys@karmanos.org (S. Eggly), slatcher@wayne.edu (R.B. Slatcher), nhagiwara@vcu.edu (N. Hagiwara).

https://doi.org/10.1016/j.socscimed.2017.12.021
Received 16 May 2017; Received in revised form 6 December 2017; Accepted 15 December 2017
Available online 18 December 2017

verbal interactions. Multiple methods are vital in that each method captures a different aspect of the interaction, and each method has its own strengths and weaknesses. Questionnaires are simple and easy to administer and capture the physician and patient perception of the interaction, but are subject to reporting biases (Bourhis et al., 1989). Observational methods may be more objective, but can involve time- and resource-intensive coding of verbal and non-verbal behaviors by at least two independent coders to assess reliability. Further, surveys and observational techniques are top-down approaches that assess only the variables of interest imposed by the researcher on the doctor-patient interaction, as well as the cultural context in which the researcher is working.

In this manuscript, we are proposing that physician and patient responsiveness can be assessed by how similar or coherent the conversation is between the physician and the patient in their interaction using a technique called Latent Semantic Analysis (LSA). LSA allows researchers to quantify the amount of semantic overlap between what a patient and a physician say to each other in a given interaction without having coders read transcripts or watch video-recorded interactions. LSA is a mathematical method for modeling semantic meaning from text (Landauer, 2007; Landauer et al., 1998). In LSA, a group of texts is processed such that each text is represented by a count of each word appearing in the text. Then, principal component analysis (a method used for dimensionality reduction) is used to derive underlying semantic dimensions. Typically semantic meaning in text can be represented using about 300 dimensions, and the meaning of a word is represented by its loading on each of the dimensions. The matrix with each word in a row and dimension loadings in each column is called the "LSA space." One notable characteristic of LSA is that the comparative meaning of two texts is not dependent on using the same words (Landauer, 2007). For example, a patient's and doctor's interaction about diabetes may be judged highly similar even if their conversation uses few overlapping words, provided the words they use have similar meanings based on the principle component analysis. Another characteristic is that it is a data-driven, "bottom-up" approach to deriving meaning, which can enable researchers to generate new theories.

LSA can quantitatively assess the semantic similarity between two texts of any length (single words, phrases, sentences, paragraphs, etc.) by correlating the dimension loadings of the word(s) in each text. Some applications of LSA have included successfully grading content adequacy of student essays (Landauer et al., 2003), diagnosing schizophrenia from patient's language as accurately as experienced psychiatrists (Elvevåg et al., 2007), and (after being trained on text similar to what an American college freshman reads) scoring as well on the Test of English as a Foreign Language as successful U.S. college applicants from non-English-speaking countries (Landauer and Dumais, 1997). Thus far LSA has been limited to assessment of semantic coherence within texts or semantic similarity across texts such as internet search applications. Only one study, to our knowledge, has applied LSA to analyze similarity in meaning between partners in a conversation. Babcock et al. (2014) employed LSA to analyze transcripts of unstructured "getting to know you" conversations between undergraduate students, and found that the semantic similarity between the interaction pairs correlated positively with behavioral ratings of number of intimate self-disclosures, verbal and nonverbal acknowledgements, number and duration of mutual gazes, and participant perception of involvement in the interaction. The authors concluded that latent semantic similarity "develops out of a highly involving interaction between mutually attentive and acknowledging partners in which a lot of verbal information is exchanged" (p. 78). These findings suggest that semantic similarity in the language used by patients and physicians as assessed by LSA can be used to assess responsiveness in the doctor-patient interaction. However, the unstructured "getting to know you" conversations between undergraduates evaluated by Babcock et al. are much different from the structured, goal-directed, and somewhat scripted communication between a physician and patient differing in status, education, and motivation during a clinical interaction.

Thus, the goal of the present paper is to demonstrate the feasibility of applying LSA methods to patient-physician communication research and assess the ability of LSA to detect similarities between what the physician and patient say in a given medical interaction. In addition, we will explore the utility of LSA in this context by presenting some preliminary results showing the sensitivity of semantic similarity to ethnic and gender variables known to affect physician-patient communication (Cooper-Patrick et al., 1999; Street et al., 2008) and using semantic similarity to predict patient outcomes of trust in the physician. Finally, we will make some methodological recommendations for using LSA to study medical interactions and suggest some research questions that can be addressed with this method.

2. Method

2.1. Participants

Participants were 132 low-income, self-identified Black/African American patients (76% women, *M*age = 43.8, *SD* = 14.0, range = 18–82) who participated in clinical interactions with 17 physicians (Mage = 27.1, range = 26–35) as part of a larger study conducted in a primary care clinic in a large midwestern city in the U.S. All physicians were second- or third-year medical residents; there were 8 from India/Pakistan (5 female) who saw 44 patients, 6 from other parts of Asia (3 female) who saw 51 patients, 2 White males who saw 33 patients, and 1 Black female who saw 4 patients. Each physician saw from 1 to 20 (median of 4) patients who participated in the study; each patient participated in only one clinical interaction. Approximately 75% of patients and 83% of physicians approached agreed to participate. For more information about participants and procedures in this study, please see the parent study from which these data were drawn (Penner et al., 2009).

2.2. Procedure

The original study was approved by the Wayne State University Behavioral IRB. The current secondary analysis of the existing de-identified transcript data was approved by the Virginia Commonwealth University IRB as an exempt study (HM14733 approved on Oct. 22, 2012). Patients completed questionnaires including demographic characteristics and previous history with medical interactions, and then participated in their medical appointment, which was video recorded. Following the interaction participants completed questionnaires about their experience in the interaction. Video recorded interactions were professionally transcribed, and transcripts were converted to raw text files and cleaned of special characters and formatting (see Hagiwara et al., 2016). All the words uttered by the patient in the interaction and all the words said by the physician in an interaction were put into separate text files, for a total of 132 patient text files and 132 physician text files.

2.3. LSA methods

In LSA each word's meaning is characterized by its loading on each of the dimensions in the semantic space, or "LSA space". The creation of a semantic space starts with a corpus of training texts, from which a word x text wordcount matrix is created. The rows of the matrix consist of each word in all of the training texts, the columns of the matrix represent each individual training text, and each cell in the matrix consists of the number of times the word in that row occurs in that column's text. The semantic space is created by performing a singular value decomposition (a form of data reduction often referred to as principal component analysis) on this word x text wordcount matrix. This yields three matrices: (1) a text x dimension matrix, giving the positions of the texts in the semantic space; (2) a word x dimension matrix, used to find the position of additional texts in the semantic space; and (3) a matrix of the "singular values" of each dimension, showing what fraction of total variance is captured by that dimension.

The first and most important choice that needs to be made in generating the semantic space is what corpus of training texts to use. The training texts can be either the texts being analyzed or an outside set of texts. While no minimum number of texts is required, more is better, and a good rule is that there should be enough texts so that there is a negligible probability of a new text adding a unique word to the corpus (Quesada, 2007). An individual text in the corpus can be of variable length, but is typically around one paragraph. A semantic space reflecting the medical meaning of words, for example, could be created by splitting a medical textbook into paragraphs and submitting each paragraph as a separate text in preparing a semantic space. In the current study, because we were unsure, given this new application of LSA, if outside corpora could capture all salient semantic features of doctor-patient interactions, we chose to create the semantic space with the texts from this study, using all the words said by each patient in an interaction and all the words said by each physician in an interaction as separate texts.

If one's data include relatively few texts, the use of a larger set of outside texts to generate the semantic space can reduce noise in inferring the semantic characteristics of the texts and allow it to capture more semantic dimensions. If outside texts are used to generate a semantic space, the choice of texts is important. LSA infers semantic meaning from the co-occurrence of words in the texts used to generate the semantic space (Quesada, 2007). Semantic spaces generated from, for example, a set of cookbooks and a set of medical textbooks will capture very different sets of semantic relationships between words. In the semantic space generated from medical textbooks, the word "sugar" will be semantically associated with words like nutrition and diabetes; in the semantic space generated from cookbooks, "sugar" will be associated with eggs, flour, and milk.

Another important choice is how many dimensions to include in the semantic space. Using too many dimensions will result in some noisy dimensions that carry no information (analogous to factors with low eigenvalues in a factor analysis), whereas if too few are included, potentially-useful information is discarded. Typically 100–300 dimensions are chosen (Martin and Berry, 2007). In the current study we used 100 dimensions because we had relatively few texts (132 physician and 132 patient texts), and the texts did not cover a wide enough range of semantic meaning (limited to physician-patient interactions), to extract a larger number of dimensions.

There are several software options available for LSA. We used the *lsa* package in R (Wild, 2015) to create our semantic space and generate the similarity coefficients analyzed in this paper. The *lsa* package is well-documented and easy to use for anyone familiar with the R programming language. However, because LSA rests on the widely-used linear algebra method of the singular value decomposition, it can be implemented with any linear algebra package, such as Python's *numpy*. If one wished to try LSA but did not have expertise in R or another available package, CU Boulder provides a web portal website (http://lsa.colorado.edu/) that allows comparison of the semantic meaning of texts by pasting texts into the web portal. This website is a valuable tool to try out LSA in that it provides clear instructions and allows easy calculation of various LSA-based metrics. Disadvantages are that it is not amenable to automation (each text comparison must be separately pasted into the web portal) and has a limited selection of semantic spaces. Additional methodological details for the current study and considerations in employing these methods are presented in a LSA tutorial available as an online-only supplement to this article.

In order to assess the semantic similarity between the physician's and patient's words in an interaction, a similarity coefficient was calculated by correlating the dimension loadings of the text from a patient in an interaction and the text by their physician in the same interaction. This correlation indicates the extent to which what a physician and a patient said to each other were semantically related. In addition, for each patient we calculated the mean of the correlations between the text of their words and all the texts of their physician talking to *other* patients during *other* examinations. Since these are correlations between the words of the patient and the words of their physician talking to other patients (e.g., in a different conversation), this correlation is expected to be near zero.

2.4. Data analysis strategy

Of interest is whether this method can detect patient-physician communication similarity, whether this measure is related to variables that might affect patient-physician communication, and whether communication similarity can predict changes in the patient's trust in the physician before and after the medical interaction. A series of regressions was conducted in order to assess whether the correlation measuring patient-physician communication similarity was significantly greater than zero, and whether it was related to physician ethnicity and gender, as well as patient gender. Then, a regression was conducted to examine whether patient trust in physicians in general prior to the interaction and patient trust in their own physician following the interaction would predict patient-physician communication similarity. When analyses involved non-independent observations (e.g., when multiple patients were seen by the same physician), we employed General Estimating Equations (GEE) regressions to correct for bias based on non-independence in the data. GEE is a form of multilevel modeling that treats group-level variation (in this case, physicians) as a random parameter and provides asymptotically normal estimates even when observations within groups are strongly correlated with one another (Hardin and Hilbe, 2003).

3. Results

The mean patient-physician communication similarity correlation was 0.142 (SD = 0.185), with a median of 0.150. This correlation was significantly greater than zero, as demonstrated by a GEE regression testing the intercept of the model, $Wald\ \chi2$ (1) = 53.96, $p < .00001$. In order to make sure that this relationship was not an artefact of conversation length or asymmetry of physician and patient speaking time, we conducted a GEE regression to analyze the relationship between the communication similarity correlations and the total length (sample mean = 2522 words; range = 565–6832) and the proportion of words spoken by the physician (mean = 0.62; range = 0.25-0.89) in each interaction. Neither conversation length ($p = .136$) nor proportion of conversation carried by the physician ($p = .843$) were related to semantic similarity. In contrast, the mean correlation between a patient's transcript and transcripts of their physician during interactions with other patients is -.005 (SD = 0.092), with a median of -.009. This mean was not different from zero, $Wald\ \chi2$ (1) = 0.469, $p < .494$. These findings indicate that LSA is specifically sensitive to the semantic relatedness between a patient's language and what their physician says while in the interaction with them, and that this similarity is not an artefact of either interaction length or conversational asymmetry.

Next, we examined whether LSA captured individual differences in patient responsiveness among physicians, and whether physicians' demographic characteristics (i.e., race/ethnicity and gender) were associated with these differences. In addition, we explored whether patient gender is associated with the semantic similarity between the patient's and physician's language in their interaction. Individual physician's mean similarity correlation with their patients ranged from 0.026 to 0.338. There was no relationship between the physician's mean similarity correlation and the number of patients the physician examined in the study, r (16) = -.21, p = .417. A one-way analysis of variance, with the 17 individual physicians as the independent variable, found that physicians differed significantly in patient-physician communication similarity in a medical interaction, F (16,115) = 1.862, $p = .031$, partial $\eta^2 = 0.206$, which is a medium to large effect size. Thus, LSA is sensitive to systematic variation across physicians in responsiveness during interactions with their patients.

A GEE regression was conducted with three physician racial/ethnic groups (White, Indian/Pakistani, other Asian) entered into the model. Because there was only one Black physician, this physician and her four patients were not included, leaving $N = 128$ for this analysis. White physicians (mean $r = 0.028$, $SE = 0.0325$) exhibited significantly lower semantic similarity with their patient's speech than Indian/Pakistani (mean $r = 0.179$, $SE = 0.024$) or other Asian physicians (mean $r = 0.185$, $SE = 0.025$), $b = -0.161$, $SE = 0.019$, *Wald* $\chi 2$ (1) = 72.82, p < .001. The latter two groups did not differ from each other. A GEE with physician gender found that female physicians had marginally greater semantic similarity with their patient's speech than did male physicians, $b = 0.065$, $SE = 0.037$, *Wald* $\chi 2$ (1) = 3.023, $p = .082$. The mean patient-physician communication similarity correlation was 0.190 for female physicians and 0.125 for male physicians. A GEE regression with patient gender entered into the model revealed that female patients' speech exhibited greater semantic similarity to their physicians than did male patients' speech, $b = 0.073$, $SE = 0.031$, *Wald* $\chi 2$ (1) = 5.72, $p = .017$. The mean patient-physician communication similarity correlation was 0.170 for female patients and 0.097 for male patients. However, there was no interaction between patient and physician gender ($p = .766$); thus gender match between patient and physician did not increase communication similarity.

A GEE regression (with physician race and gender, patient gender, and pre-interaction trust covaried) found that greater communication similarity was associated with *less* trust in physicians in general reported by the patient prior to the interaction, $b = -0.027$, SE = 0.0087, Wald $\chi 2$ (1) = 9.591, $p = .002$, and by *greater* patient trust in their own physician following the interaction, $b = 0.032$, $SE = 0.0123$, *Wald* $\chi 2$ (1) = 6.703, $p < .010$. Neither pre- nor post-interaction trust was significantly associated with communication similarity when analyzed separately. It should be noted that, because the trust questionnaire was not added until later in the study, the sample size was smaller ($N = 65$ with 15 different physicians) for this analysis.

4. Discussion

The current study is the second study we know of to apply Latent Semantic Analysis (LSA), a technique that assesses the latent semantic meaning in language, to analyze a conversation, and the first to evaluate the semantic similarity in a medical interaction. There were several important findings. First, LSA was sensitive to similarities between what the physician and patient said in a given medical interaction: There was a significant positive relationship between the speech of the patient and physician in their interaction, whereas no relationship was found between the patient's words in one interaction and the same physician's words from interactions with different patients. Further, physicians exhibited reliable differences between each other in the semantic similarity of their language with the language of their patients, suggesting LSA is sensitive to stable individual differences in the way physicians interact with their patients.

The relationship between semantic similarity and three physician and patient characteristics were explored: physician race/ethnicity, physician gender, and patient gender. The present study demonstrated that White physicians' conversations exhibited lower semantic similarity with their patients than did physicians with an Indian/Pakistani or other Asian background, that interactions with female physicians exhibited marginally greater semantic similarity with their patients than did their male counterparts, and that interactions with female patients displayed greater semantic similarity than interactions involving male patients. Some exploratory findings also suggested that physician-patient communication similarity is related to patient expectations of and perceptions about the interaction: greater generalized trust of doctors prior to the interaction was associated with *less* communication similarity, and greater specific trust in their own physician following the interaction was associated with *more* communication similarity. This intriguing though preliminary result may indicate that a patient entering a medical interaction with greater generalized trust in physicians is less motivated to question their physician or assert their needs and opinions in the interaction (Trachtenberg et al., 2005), whereas an interaction with a physician that is characterized by communication similarity is associated with greater subsequent trust in that physician.

Caution is needed in generalizing these preliminary findings, especially with regard to the analyses of physician demographic characteristics. There were only one Black and two White physicians representing these racial/ethnic groups, greatly reducing generalizability. Because of the unbalanced distribution of physician race/ethnicity and gender (e.g., both White physicians were males and the one Black physician was a female) the physician race/ethnicity and gender effects could not be examined independently, and these effects may be confounded. Similarly, because all patients in the study self-identified as Black/African American, it is not possible to determine whether the effects reported here are due to general communication styles or are specific to interactions with Black/African American patients. Further, the overall total sample size ($N = 132$) is small, and because the measure of patient trust was added midway through the study sample size is further reduced for this analysis. Nevertheless, these initial findings encourage further use of LSA to investigate communication patterns in patient-physician interaction. Future research should use larger datasets to investigate questions about the effects of demographic influences, individual differences, and contextual factors on patient-physician communications, and to examine the association of semantic similarity with important variables such as patient satisfaction, trust, adherence, and health outcomes.

Several considerations should be kept in mind when deciding whether to apply LSA to physician-patient communication research. First, although LSA requires fewer resources to evaluate text when compared to human coding of medical interactions, the research team still needs to record and transcribe the medical interactions and format the texts for computer analysis. Second, interpretation of results is dependent on the semantic space used in the study, and so generalization across studies needs to be done with caution; this limitation may be obviated by sharing semantic spaces across laboratories. Third, in LSA meaning is derived from the co-occurrence of words in each text, without regard to the order of words, punctuation, or nearness of words within the text. However, the amount of meaning lost by not considering word order may be small; several methods converge to find that word order accounts for about 10–15% of the variance in the meaning of multi-sentence English texts (Landauer, 2007). Finally, because LSA is a bottom-up, atheoretical approach, it is difficult to derive specific semantic meaning from dimension loadings or the semantic space. However, this characteristic of LSA may also provide some advantages over investigator-created coding scales and questionnaires, by allowing discovery of connections that exist outside of the theoretical and cultural constructs that may currently constrain investigators. The online supplement presents more detailed considerations when using LSA, as well as other potential applications of the method, such as investigating interactions with more granularity and studying patient-physician-companion triads. LSA flexibly allows assessment of similarity between texts of any length, from single words to lengthy manuscripts, and permits asking a variety of research questions about face-to-face or tele-health interactions, or health-related internet sites (e.g., Jucks and Bromme, 2007). In sum, LSA is a very promising tool for patient-physician communication researchers.

Acknowledgements

Videorecorded interactions were created during a study supported by NICHD R21HD050445001 awarded to the third author, and transcript creation was supported by NINR R03NR013249 awarded to the sixth author.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.socscimed.2017.12.021.

References

Babcock, M.J., Ta, V.P., Ickes, W., 2014. Latent semantic similarity and language style matching in initial dyadic interactions. J. Lang. Soc. Psychol. 33, 78–88.
Beach, M.C., Inui, T., the Relationship-Centered Care Research Network, 2006. Relationship-centered care: a constructive reframing. J. Gen. Intern. Med. 21 (S1), S3–S8.
Boon, H., Stewart, M., 1998. Patient-physician communication assessment instruments: 1986 to 1996 in review. Patient Educ. Counsel. 35, 161–176. http://dx.doi.org/10.1016/S0738-3991(98)00063-9.
Bourhis, R.Y., Roth, S., MacQueen, G., 1989. Communication in the hospital setting: a survey of medical and everyday language use amongst patients, nurses and doctors. Soc. Sci. Med. 28, 339–346 org.proxy.library.vcu.edu/10.1016/0277-9536(89)90035-X.
Cooper-Patrick, L., Gallo, J.J., Gonzales, J.J., Vu, H.T., Powe, N.R., ... Ford, D.E., 1999. Race, gender, and partnership in the patient-physician relationship. J. Am. Med. Assoc. 282 (6), 583–589. http://dx.doi.org/10.1001/jama.282.6.583.
Davis, K., Schoenbaum, S.C., Audet, A., 2005. A 2020 vision of patient-centered primary care. J. Gen. Intern. Med. 20 (10), 953–957.
Elvevåg, B., Foltz, P.W., Weinberger, D.R., Goldberg, T.E., 2007. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. Schizophr. Res. 93, 304–316.
Epstein, R.M., Franks, P., Fiscella, K., Shields, C.G., Meldrum, S.C., Kravitz, R.L., Duberstein, P.R., 2005. Measuring patient-centered communication in patient-physician consultations: theoretical and practical issues. Soc. Sci. Med. 61 (7), 1516–1528.
Epstein, R.M., Street, R.L., 2007. Patient-centered Communication in Cancer Care: Promoting Healing and Reducing Suffering. National Cancer Institute, Bethesda, MD.
Epstein, R.M., Street, R.L., 2011. The values and value of patient-centered care. Ann. Fam. Med. 9 (2), 100–103.
Hardin, J.W., Hilbe, J.M., 2003. Generalized Estimating Equations. Chapman & Hall/CRC, Boca Raton, FL.
Hagiwara, N., Slatcher, R.B., Eggly, S., Penner, L.A., 2016. Physician racial bias and word use during racially discordant medical interactions. Health Commun. 32, 401–408.
Institute of Medicine, 2001. Crossing the Quality Chasm: a New Health System for the Twenty-first Century. National Academies Press, Washington DC.
Jucks, R., Bromme, R., 2007. Choice of words in doctor-patient communication: an analysis of health-related internet sites. Health Commun. 21, 267–277.
Landauer, T.K., 2007. LSA as a theory of meaning. In: Landauer, T.K., McNamara, D.S., Dennis, S., Kintsch, W. (Eds.), Handbook of Latent Semantic Analysis. Lawrence Erlbaum Associates, Inc., Publishers, Mahwah, NJ, pp. 3–34.
Landauer, T.K., Dumais, S.T., 1997. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. Psychol. Rev. 104, 211.
Landauer, T.K., Laham, D., Foltz, P., 2003. Automatic essay assessment. Assess Educ. Princ. Pol. Pract. 10, 295–308.
Landauer, T.K., Foltz, P.W., Laham, D., 1998. An introduction to latent semantic analysis. Discourse Process 25, 259–284.
Martin, D.I., Berry, M.W., 2007. Mathematical foundations behind latent semantic analysis. In: Landauer, T.K., McNamara, D.S., Dennis, S., Kintsch, W. (Eds.), Handbook of Latent Semantic Analysis. Lawrence Erlbaum Associates, Inc., Publishers, Mahwah, NJ, pp. 35–55.
Matusitz, J., Spear, J., 2014. Effective doctor-patient communication: an updated examination. Soc. Work. Publ. Health 29 (3), 252–266. http://dx.doi.org/10.1080/19371918.2013.776416.
Ong, L.M.L., De Haes, J.C.J.M., Hoos, A.M., Lammes, F.B., 1995. Doctor-patient communication: a review of the literature. Soc. Sci. Med. 40 (7), 903–918.
Penner, L.A., Dovidio, J.F., Edmondson, D., Dailey, R.K., Markova, T., Albrecht, T.L., Gaertner, S.L., 2009. The experience of discrimination and Black-White health disparities in medical care. J. Black Psychol. 35, 180–203.
Quesada, J., 2007. Creating your own LSA spaces. In: Landauer, T.K., McNamara, D.S., Dennis, S., Kintsch, W. (Eds.), Handbook of Latent Semantic Analysis. Lawrence Erlbaum Associates, Inc., Publishers, Mahwah, NJ, pp. 71–85.
Roter, D., 2000. The enduring and evolving nature of the patient-physician relationship. Patient Educ. Counsel. 39 (1), 5–15.
Stewart, M.A., 1995. Effective physician-patient communication and health outcomes: a review. Can. Med. Assoc. J. 152 (9), 1423–1433.
Stewart, M.A., Brown, J.B., Weston, W.W., McWhinney, I.R., McWilliam, C.L., Freeman, T.R., 2003. Patient-centered Medicine: Transforming the Clinical Method, second ed. Radcliffe Medical Press, Oxon, United Kingdom.
Street Jr., R.L., O'Malley, K.J., Cooper, L.A., Haidet, P., 2008. Understanding concordance in patient-physician relationships: personal and ethnic dimensions of shared identity. Ann. Fam. Med. 6 (3), 198–205. http://dx.doi.org/10.1370/afm.821.
Suchman, A.L., 2006. A new theoretical foundation for relationship-centered care. J. Gen. Intern. Med. 21 (S1), S40–S44.
Trachtenberg, F., Dugan, E., Hall, M.A., 2005. How patients' trust relates to their involvement in medical care. J. Fam. Pract. 54 (4), 344e352.
Wild, F., 2015. lsa. Retrieved from. https://CRAN.R-project.org/package=lsa.