ARTICLE

Buildup of speaking skills in an online learning community: a network-analytic exploration

Rasoul Shafipour[1], Raiyan Abdul Baten[1], Md Kamrul Hasan[2], Gourab Ghoshal[3], Gonzalo Mateos[1] & Mohammed Ehsan Hoque[2]

ABSTRACT   Studies in learning communities have consistently found evidence that peer-interactions contribute to students' performance outcomes. A particularly important competence in the modern context is the ability to communicate ideas effectively. One metric of this is speaking, which is an important skill in professional and casual settings. In this study, we explore peer-interaction effects in online networks on speaking skill development. In particular, we present an evidence for gradual buildup of skills in a small-group setting that has not been reported in the literature. Evaluating the development of such skills requires studying objective evidence, for which purpose, we introduce a novel dataset of six online communities consisting of 158 participants focusing on improving their speaking skills. They video-record speeches for 5 prompts in 10 days and exchange comments and performance-ratings with their peers. We ask (i) whether the participants' ratings are affected by their interaction patterns with peers, and (ii) whether there is any gradual buildup of speaking skills in the communities towards homogeneity. To analyze the data, we employ tools from the emerging field of Graph Signal Processing (GSP). GSP enjoys a distinction from Social Network Analysis in that the latter is concerned primarily with the connection structures of graphs, while the former studies signals on top of graphs. We study the performance ratings of the participants as *graph signals* atop underlying interaction topologies. Total variation analysis of the graph signals show that the participants' rating differences decrease with time (slope = -0.04, $p < 0.01$), while average ratings increase (slope = 0.07, $p < 0.05$)—thereby gradually building up the ratings towards community-wide homogeneity. We provide evidence for peer-influence through a prediction formulation. Our consensus-based prediction model outperforms baseline network-agnostic regression models by about 23% in predicting performance ratings. This in turn shows that participants' ratings are affected by their peers' ratings and the associated interaction patterns, corroborating previous findings. Then, we formulate a consensus-based diffusion model that captures these observations of peer-influence from our analyses. We anticipate that this study will open up future avenues for a broader exploration of peer-influenced skill development mechanisms, and potentially help design innovative interventions in small-groups to maximize peer-effects.

[1] Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY, USA. [2] Department of Computer Science, University of Rochester, Rochester, NY, USA. [3] Department of Physics and Astronomy, University of Rochester, Rochester, NY, USA. These authors contributed equally: Rasoul Shafipour, Raiyan Abdul Baten. Correspondence and requests for materials should be addressed to R.S. (email: rshafipo@ur.rochester.edu)

Introduction

The idea that learners actively construct knowledge as they build mental frameworks to make sense of their environments has been widely studied under the *cognitivism* and *constructivism* theories of learning (Jones, 2015; Harasim, 2017). For example, in a research on cognitive psychology, subjects were shown the sentence, *The window is not closed.* Later, most of them recalled the sentence as *The window is open* (Cross, 1998), indicating that people build a mental image or a cognitive map, a *schema*, in the learning process, and later build new knowledge in connection to what they already know (Mayes and De Freitas, 2013; Harasim, 2017).

Peer interaction is established as an important way people learn (Topping et al., 2017; Johnson and Johnson, 2009). In the process of collaboration, learning occurs as individuals build and improve their mental models through discussion and information sharing with their peers in small groups (Davidson and Major, 2014; Leidner and Jarvenpaa, 1995). Research on learning communities has investigated how social and academic networks contribute to students' learning experiences and outcomes (Brouwer et al., 2017; Smith, 2015; Smith et al., 2009; Celant, 2013). In particular, study-related, task-dependent academic support relationships have been shown to lead to significantly higher academic performance (Gašević et al., 2013; Frank et al., 2008). In soft-skill development, Fidalgo-Blanco et al., (2015) explored peer-interaction effects in developing teamwork competency in small groups, and found a direct relation between quantities of interaction and individual performance.

In this paper, we explore the scope of peer-induced *speaking skill* development, namely, the influence of peer interactions in small groups towards improving one's speaking competencies. The ability to communicate ideas and thoughts effectively is valued to be an important skill throughout professional and casual interaction settings. Given this, can individuals improve their speaking skills by interacting with peers in small groups? If so, can such effects be extracted from data objectively, and then correctly modeled? To the best of our knowledge, there have been limited studies conducted in the past in this particular context. Part of the problem is in identifying or indeed generating relevant datasets. In particular, the monitoring and evaluation of peer effects objectively require data with quantifiable evidence. However, keeping a structured history of peer interactions had proved difficult in traditional offline studies, leading to the use of questionnaires and self-reported data (Lomi et al., 2011; Huitsing et al., 2012).

With the recent advent of online communication, it is now commonplace to track detailed specifics of the participants' interactions from their online footprint, such as from Massive Open Online Courses' (MOOC) discussion forums (Tawfik et al., 2017; Brinton et al., 2014). Online learning communities and other computer supported collaborative learning platforms thus facilitate objective studies of learner interaction and associated outcomes (Jones, 2015; O'Malley, 2012; Dado and Bodemer, 2017; Russo and Koesten, 2005; Palonen and Hakkarainen, 2013). Consequently, our first contribution in this paper is in constructing a novel dataset of six online communities, where the participants focus on developing their *speaking* skills. The 158 participants in the six communities record video speeches in response to common job interview prompts, and subsequently exchange feedback comments and performance ratings with peers in their respective communities. They respond to 5 prompts across 10 days, thereby generating a unique *temporal* dataset of comment-based interactions and speaking performance ratings. In the dataset, the participants leave a comment only after they have watched a video of a peer. In such a pairwise interaction, Participant A can potentially learn from the feedback given by

Participant B, and Participant B can also learn from Participant A's public speaking approach, good practices, and mistakes by watching the video. Moreover, the meta-cognitive task of explaining something through feedback can also clarify Participant B's own understanding (De Backer et al., 2015; Roscoe and Chi, 2008). In line with the collaborative model of peer-induced learning, one might expect these interactions to help the participants gradually build mental models of the good practices of speaking (Davidson and Major, 2014). Knowledge regarding speaking skills has a tacit nature, which in turn can be developed through observation, imitation, and practice (Busch, 2008; Chugh et al., 2013). Furthermore, knowledge is a non-rival good and can be traded without decreasing the level possessed by each trader. Therefore, such mutually beneficial interactions can be anticipated to lead to a gradual accumulation of speaking skills in the communities—as measured by the performance ratings—eventually driving the communities towards skill homogeneity as time goes on. Whether or not such gradual buildup of skills happens in a small group setting has not been reported in learning community literature, and that is a gap we address in this study. The temporal nature of our dataset facilitates such exploration.

Interactions in a learning platform are well suited to be modeled as a complex network, where the learners are denoted as nodes and their interactions as edges. This naturally leads to the use of Social Network Analysis (SNA) tools to gain insights on how various network parameters affect networked learning. For example, network position parameters (closeness and degree centrality, prestige etc.) have been shown to influence final grades (Putnik et al., 2016; Cho et al., 2007) and cognitive learning outcomes (Russo and Koesten, 2005) of the learners. However, SNA is concerned with the *structure of relational ties* between a group of actors (Carolan 2013), and not with any intrinsic measures of 'performance' or 'merits' of those actors. For example, if a person has a lot of followers on Twitter, the in-degree measure will be high, and one would say the person has a high in-degree centrality (in lay terms, enjoys a prominent position in the network) (Newman, 2010). This intuition of 'importance' is extracted from the connection topology, not from any intrinsic measure of how good the person's tweets actually are. Therefore, in a learning setting, researchers usually apply correlation analysis and other statistical methods to investigate the relations between SNA measures and performance/learning outcomes (Dado and Bodemer, 2017). Recent developments in the emerging field of Graph Signal Processing (GSP) provide a novel way to integrate such 'performance' measures (signals) with network graphs (Shuman et al., 2013). The idea is to attach a signal value to every node in the graph, and process the signals atop the underlying graph structure. For example, Huang et al., (2016) used GSP to study brain imaging data, where the brain structure is modeled as a graph and the brain activities as signals. Deri and Moura (2016) exploited GSP to study vehicle trajectories over road networks. Other recent application domains include Neuroscience (Rui et al., 2017), imaging (Pang and Cheung, 2016; Thanou et al., 2016), medical imaging (Kotzagiannidis and Dragotti, 2016), to name a few. In a learning setting, these developments in GSP open up the opportunity to model learners' academic grades or similar outcome measures as *graph signals* on top of a peer interaction network. Consequently, our second contribution in this paper is in mining evidences of peer-influence and buildup of speaking skills using novel Graph Signal Processing based techniques. More formally, we ask (1) whether the learners' ratings are affected by their interaction patterns and the ratings of the peers they interact with, and (2) whether there is any gradual buildup of speaking skills in the communities towards homogeneity.

Towards answering the aforementioned research questions, we take two approaches. *First*, we measure the smoothness of the graph signals (speaking ratings) as the participants interact temporally across 5 prompts. We show that with time, people's ratings come closer to the ratings of the peers they interact with and the ratings also increase on average. This suggests that the communities in the dataset gradually approach homogeneity in terms of the performance ratings of the interacting participants. *Second*, we approach the mining of peer-influence as a prediction formulation. Let us introduce the idea through a stylized scenario. In a classroom setting, a teacher can track the test scores of a student and make a prediction of his/her future score by running a linear regression through previous scores, providing a baseline trend. Instead, if the teacher also takes into account the scores of the student's peers (those that the student interacts with), we demonstrate that the prediction of future scores for both the student and the peers outperform the baseline. More formally, a network-consensus prediction model outperforms a network-agnostic model, a trend we demonstrate to be true in all six communities in our generated data. This corroborates previous findings that interaction quantities and peers' grades/performance outcomes impact learners' own performance outcomes (Fidalgo-Blanco et al., 2015; Hoxby, 2000), and we show that they hold for speaking skill development as well.

We further proceed to model and simulate the peer-effects observed in the dataset. From an economic lens, various models for knowledge, information or opinion flow across a network has been explored in previous literature (Golub and Sadler, 2017; Cowan and Jonard, 2004; Gale and Kariv, 2003). We introduce a model of peer-induced knowledge propagation (in terms of speaking performance ratings) based on consensus protocols for network synchronization and distributed decision making (Olfati-Saber et al., 2007). Olfati-Saber and Murray (2004) studied such consensus algorithms and presented relevant convergence analysis. We modify their protocol to suit the characteristics of the communities of our dataset. We discuss how the proposed diffusion dynamics capture the peer-influence and the gradual buildup of performance ratings towards community-wide homogeneity, as observed in our dataset.

In this context, our contributions can be summarized as follows:

- Constructing a dataset of six online communities that allow studying longitudinal peer-effects in speaking skill development;
- Yielding new evidences for skills gradually building up in learning communities towards homogeneity, corroborating previous findings of positive impacts of peer-influence, and modeling the observations via diffusion dynamics;
- Employing novel Graph Signal Processing tools for extracting the insights.

Dataset
Data collection. The data comes from a ubiquitous online system, ROC Speak (Fung et al., 2015), that gives people semi-automated feedback on public speaking. One hundred and fifty eight participants, aged 18 to 54 years, were hired from Amazon Mechanical Turk. All of them were native English speakers, and came from a variety of professional and educational backgrounds. The participants were randomly assigned into 6 groups labeled 1 through 6, which had 26, 31, 26, 30, 22, and 23 participants respectively. They were assigned a common goal of improving their speaking skills, towards which they were given 5 common job interview prompts in 10 days. Responses to these prompts were recorded via webcam as videos (typically 2 min in length)

and stored in the platform. The study was conducted upon approval from the authors' University IRB. Note, that in what is to follow, we use the terms groups and communities interchangeably.

Automated feedback. For groups 2, 4, 6, the system generated automatic feedback on smile intensity, body movement/gesture, loudness, pitch, unique word count, word cloud and instances of weak language use, as shown in Fig. 1. Groups 1, 3, 5 did not receive any automated feedback.

Feedback comments and ratings. In all six groups, the participants exchanged feedback with their peers in their respective groups. They were required to give feedback to at least three peers in each prompt, whom they could choose at will from a feed of all their peers' videos. Each feedback comprised of (1) at least three comments and (2) performance ratings on a 1–5 Likert scale.

The participants were not given any explicit instruction on what aspects of speaking skills to give comments on. However, the user interface allowed them to tag the comments whether they were on friendliness, volume modulation or use of gestures. The comments were mostly focused on speech delivery. Here, we do not concern ourselves with an analysis of the specific contents of the comments, instead refer to each comment as an 'interaction' unit.

When uploading a video, participants could select 5 qualities from a list of 23 qualities on which they wanted performance ratings from their peers. For example, the video in Fig. 1 asks for ratings on the qualities of Attention Grabbing, Explanation of Concepts, Credibility, Vocal Emphasis, Appropriate Pausing, as shown in segment D of the figure. In addition to these customized rating categories, all videos received 'overall' delivery performance ratings from the peers, as shown in segment D of Fig. 1. The participants were at liberty to use their judgments in giving ratings to their peers. For our analysis, we use the average 'overall' rating each video received from the peers and refer to this average value as 'peer rating', 'performance rating' or 'speaking rating' interchangeably in the sequel, considering it to be an objective abstraction of the participant's overall performance or skill level in delivering the speech.

Dataset summary. All of the 158 participants in the 6 groups completed the study by recording at least one video to all 5 prompts—generating a total of 817 videos. The 6 groups have 25,665 peer-generated ratings in total. Out of them, 5053 are 'overall' performance ratings that we use in this study. Therefore, each video received 'overall' ratings from 6.19 peers on average. The dataset has 14,285 comments in total, with an average of 17.49 comments per video.

Methodology
In this section, we detail the methods of mining peer-influence evidence in the dataset. As argued in the Introduction, we are interested to see if the interactions impact the development of speaking-related skills in the six groups. To reiterate, we examine (1) whether the learners' ratings are affected by their interaction patterns and the ratings of the peers they interact with, and (2) whether there is any gradual buildup of speaking skills in the communities towards homogeneity.

Performance ratings as graph signals. Consider modeling an online community's participants and their interactions through a network graph. To that end, we naturally identify each participant with a node in the said graph. As discussed in the Introduction, each comment-based interaction can potentially influence both
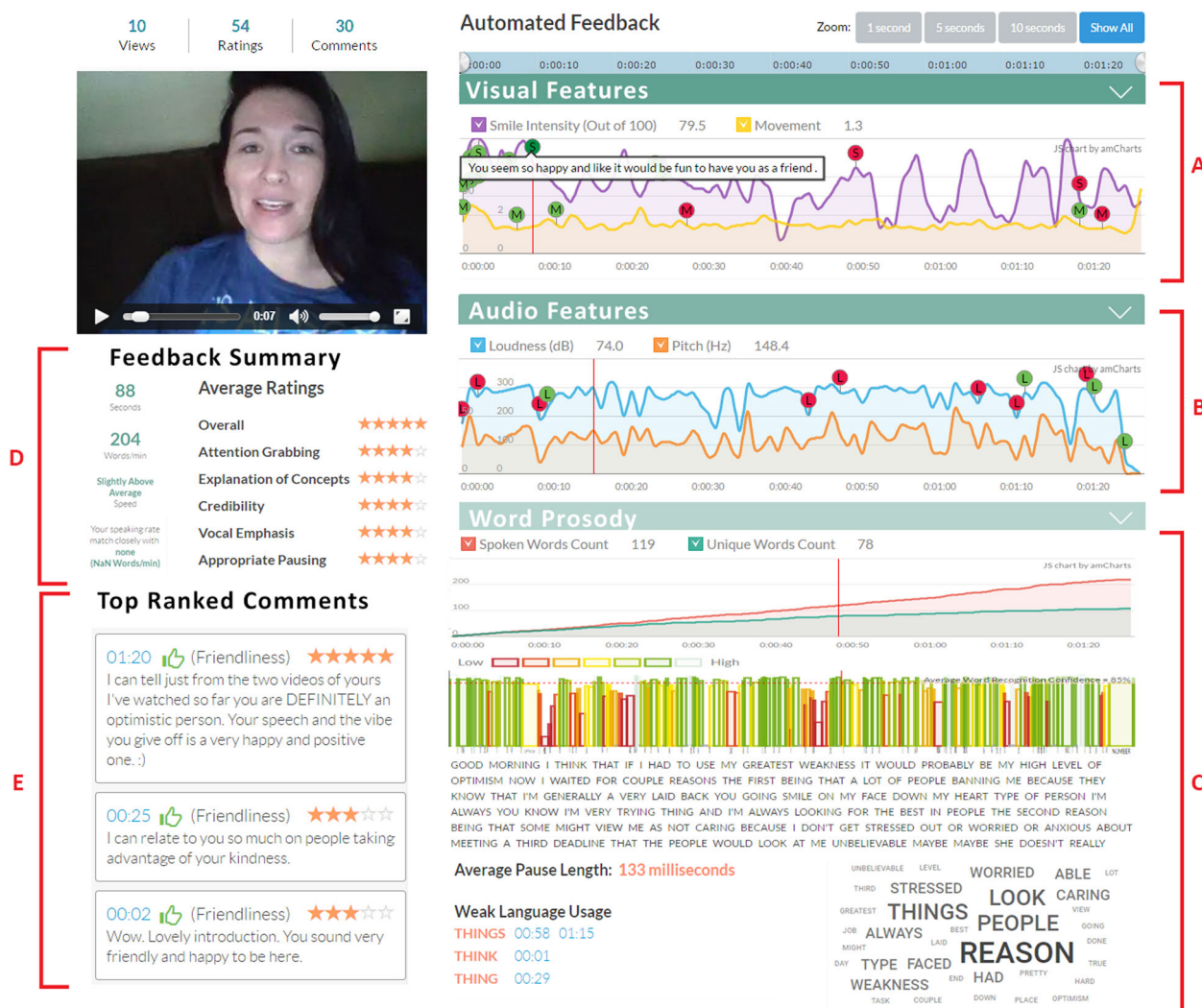
Fig. 1 A snapshot of the ROC Speak feedback interface. The page shows automatically generated measurements for (**a**) smile intensity, gestural movements, (**b**) loudness, pitch, (**c**) unique word count, word recognition confidence, a transcription of the speech, word cloud, and instances of weak language. Additionally, the peer-generated feedback are shown in (**d**) a Feedback Summary section, and (**e**) a top ranked comments section classified by usefulness and sentiment. In our study, we use the overall ratings given by the peers, as shown in segment (**d**)

the commenter and the receiver. The receiver can benefit from the comment itself. On the other hand, the commenter can learn from watching the peer's video and from the meta-cognitive task of providing a feedback. Therefore, each interaction is represented via an undirected edge, acknowledging mutual benefit for the two nodes. Exchanges of multiple comments are encoded through integer-valued edge weights; more formally, consider an undirected, weighted graph $G(\mathcal{N}, \mathcal{E}, \mathbf{W}^{(p)})$ representing this network at pth prompt with a node set $\mathcal{N}$ of known cardinality N (i.e., the total number of participants in the group—each of the groups in the dataset have their own cardinality), and the edge set $\mathcal{E}$ of unordered pairs of elements in $\mathcal{N}$. The so-called symmetric weighted adjacency matrix is denoted by $\mathbf{W}^{(p)} \in \mathbb{R}^{N \times N}$, whose ijth element represents the total number of comments that participants i and j have given to each other in the pth prompt. Figure 2 illustrates the construction of $\mathbf{W}$ using a toy example. Since the interaction patterns tend to change from prompt to prompt, so does the connectivity pattern (i.e., the topology) of the resulting graph and hence the explicit dependency of the weights in $\mathbf{W}^{(p)}$ with respect to p.

Next, we incorporate the participants' rating information in the form of graph signals. Here, a *graph signal* (Shuman et al., 2013)

is a vertex-valued network process that can be represented as a vector of size N supported on the nodes of G, where its ith component is the rating of node i. As explained in the Dataset section, the participants are given overall performance ratings by their peers, and we take a participant's average overall rating in each prompt as the graph signal value. Thus, we collect the ratings of the pth prompt in a vector $\mathbf{r}^{(p)} \in \mathbb{R}^N$, where $r_i^{(p)}$ is ith participant's rating. This representation is visualized in Fig. 2. Under the natural assumption that signal properties are related to graph topology, the goal of Graph Signal Processing is to develop models and algorithms that fruitfully leverage this relational structure, making inferences about these signal values when they are only partially observed (Shuman et al., 2013).

Total variation analysis. We explore whether the communities gradually approach homogeneity in performance ratings, by conducting a total variation analysis. Total variation is a measure of smoothness of the graph signals (ratings), the idea being, if non-rival interactions allow participants to gradually achieve ratings closer to their peers, then ratings will become smoother or more homogeneous across the network.
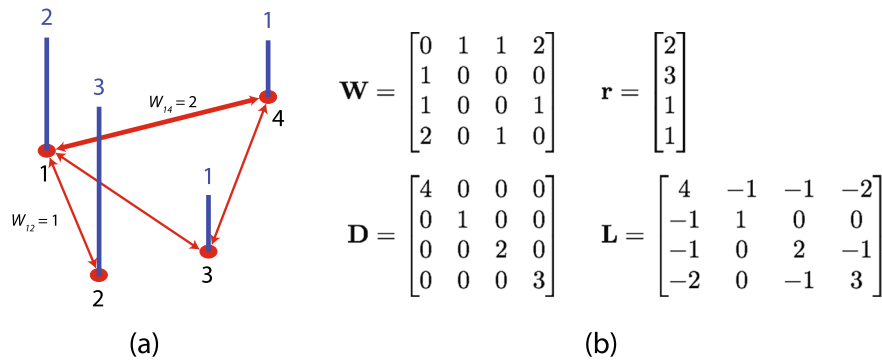
Fig. 2 a A network graph representation of an example community with $N = 4$ participants depicted as red nodes. The edges denote interactions in the form of feedback comments, with the edge weights corresponding to the number of interactions between nodes. As indicated by edge weights, participants 1 and 4 interact 2 times, while all other nonzero pairwise interactions take place only once. The blue bars on top of the nodes represent average speaking ratings as *graph signals*, and can take any value between 1 and 5. **b** The Adjacency matrix **W** captures the number of interactions between participants i and j in its ijth element. Graph signal **r** captures everyone's speaking performance ratings. The diagonal degree matrix **D** has node i's number of interactions in its iith entry. The Laplacian L is computed by **D**−**W**. In our dataset, **W**, **r**, **D**, and **L** vary across prompts p, and also across the six groups

Given the network adjacency matrix $\mathbf{W}^{(p)} \in \mathbb{R}^{N \times N}$, the degree (number of links) of participant i at the pth prompt is the total number of comments exchanged in that prompt, defined as $d_i^{(p)} := \sum_j W_{ij}^{(p)}$. Using these definitions, we construct the Laplacian matrix for any prompt p according to $\mathbf{L}^{(p)} := \mathbf{D}^{(p)} - \mathbf{W}^{(p)}$, where $\mathbf{D}^{(p)} := \mathrm{diag}(\mathbf{W}^{(p)}\mathbf{1}_N)$ is a diagonal matrix with elements corresponding to the nodes' degree, and $\mathbf{1}_N$ are the vector of ones of length N (Chung, 1997). Figure 2 illustrates the construction of the degree and Laplacian matrices. For a graph signal **r**, one can utilize the graph Laplacian to compute the total variation (TV) of the participants' ratings thus,

$$\mathrm{TV}(\mathbf{r}) := \mathbf{r}^T \mathbf{L} \mathbf{r} = \sum_{i,j=1, i>j}^{N} W_{ij}\left(r_i - r_j\right)^2, \quad (1)$$

where r_i and r_j reflect the ratings of participants i and j, while weights W_{ij} account for their volume of interaction. The total variation in Eq. (1) is often referred to as a smoothness measure of the signal **r** with respect to the graph G. If two nodes i and j do not interact in a prompt, W_{ij} is 0 and therefore $(r_i - r_j)^2$ makes no contribution to the sum in Eq. (1), while increasing interaction leads to more contributions from the term $(r_i - r_j)^2$. If **r** differs significantly between pairs of nodes that show strong patterns of interaction, then TV(**r**) is expected to be large. On the other hand, if ratings vary negligibly across connected nodes, then the graph signal is smooth and TV(**r**) takes on small values. Increasing convergence of ratings across interacting participants leads to a decreasing total variation. However, TV does not capture the overall trend of the ratings—whether the participants' ratings are improving or not. Therefore, to complement the total variation analysis, we calculate the network-wide average ratings at each prompt as

$$\bar{r}^{(p)} := \frac{1}{N}\sum_{i=1}^{N} r_i^{(p)}, \quad (2)$$

where $r_i^{(p)}$ is the rating of the ith participant in the pth prompt.

Prediction under smoothness prior. In the Introduction, we presented a scenario of a teacher tracking the performance rating evolution of students to illustrate our approach of prediction formulation towards shedding light on peer-effects. More formally, we formulate a convex optimization problem to predict future performance ratings. We compare prediction errors between two models: (i) an ordinary least-squares regression model using only

past performance ratings of the participant, i.e., a model which is agnostic to the network effects on the participant; and (ii) augmenting the least-squares predictor with a smoothness regularization term encouraging network-wide consensus. In other words, the second model considers the ratings of one's peers and the quantities of interaction along with one's own records. It is important to note that we do not seek to fit the data to find the best possible regression model for predicting future ratings, rather, we intend to compare prediction errors between network-agnostic and consensus-based regression models to illustrate our point that the network has an impact on individual outcomes.

We use the first four prompts in our data for training purposes, and predict the 5th prompt's ratings for all the participants in the six groups. Specifically, we collect the rating signal vectors $\mathbf{r}^{(p)} \in \mathbb{R}^N$ for prompts $p = 1, 2, 3, 4$ to form the training set and predict the ratings at prompt $p = 5$ using two different linear regression models:

$$\hat{\mathbf{r}} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p \quad (3)$$

and

$$\hat{\mathbf{r}} = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p + \boldsymbol{\beta}_2 \sqrt{p}, \quad (4)$$

where $\hat{\mathbf{r}} \in \mathbb{R}^N$ is the estimated ratings at pth (in our case, 5th) prompt. For Eqs. (3) and (4), the parameters $\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2 \in \mathbb{R}^N$ are learned through the following regularized least-squares criteria, respectively:

$$\{\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\beta}}_1\} = \underset{\boldsymbol{\beta}_0, \boldsymbol{\beta}_1}{\mathrm{argmin}}\frac{1}{2m}\sum_{p=1}^{m}\left\|\mathbf{r}^{(p)} - (\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p)\right\|^2$$
$$+ \lambda \sum_{p=1}^{m}(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p)^T \mathbf{L}^{(p)}(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p) \quad (5)$$

and

$$\{\hat{\boldsymbol{\beta}}_0, \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2\} = \underset{\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2}{\mathrm{argmin}}\frac{1}{2m}\sum_{p=1}^{m}\left\|\mathbf{r}^{(p)} - (\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p + \boldsymbol{\beta}_2 \sqrt{p})\right\|^2$$
$$+ \lambda \sum_{p=1}^{m}(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p + \boldsymbol{\beta}_2 \sqrt{p})^T$$
$$\times \mathbf{L}^{(p)}(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 p + \boldsymbol{\beta}_2 \sqrt{p})$$
$$+ \mu\left(\|\boldsymbol{\beta}_1\|^2 + \|\boldsymbol{\beta}_2\|^2\right). \quad (6)$$

Here,c non-linear $\sqrt{p}$ term to capture any saturation effect at limiting time. We have $m = 4$ since the training is done on the first four prompts' data. Let us break down the objective functions Eqs. (5) and (6) for clarity. The first summands in Eqs. (5) and (6) are data fidelity terms, which take into account each participant's individual trajectory of learning across the first m prompts, and fits them to the respective linear regression models. Accordingly, the first terms can be attributed to a person's own talent or pace of learning, agnostic to the network. The second summands in Eqs. (5) and (6) incorporate the effects of the neighbors' ratings and the amount of interaction into a participant's learning curve. Notice that these terms are nothing else than smoothing regularizers of the form $\sum_{p=1}^{m} \mathrm{TV}(\widehat{\mathbf{r}})$, hence encouraging participant ratings prediction with small total variation. The tuning parameter $\lambda > 0$ balances the trade-off between faithfulness to the past ratings data and the smoothness (in a total variation sense) of the predicted graph signals $\widehat{\mathbf{r}} = \widehat{\boldsymbol{\beta}}_0 + \widehat{\boldsymbol{\beta}}_1 p$ and $\widehat{\mathbf{r}} = \widehat{\boldsymbol{\beta}}_0 + \widehat{\boldsymbol{\beta}}_1 p + \widehat{\boldsymbol{\beta}}_2 \sqrt{p}$, and can be chosen via model selection techniques such as cross validation (Friedman et al., 2001). Parameter μ which is chosen along with λ via leave-one-out cross validation, prevents overfitting via a shrinkage mechanism which ensures that none of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ get to dominate the objective functions disproportionately (Friedman et al., 2001). Equations (5) and (6) are both convex, specifically unconstrained quadratic programs that can be solved efficiently (Boyd and Vandenberghe, 2004) via off-the-shelf software packages. Here we use CVX (Grant et al., 2008), a MATLAB-based modeling system for solving convex optimization problems. An outline of the procedure is presented in Algorithm 1.

Algorithm 1 Prediction Under Smoothness Prior

Input: Rating vectors $\{\mathbf{r}^{(p)}\}_{p=1}^m$ and Laplacian matrices $\{\mathbf{L}^{(p)}\}_{p=1}^m$
Output: Predicted ratings at $p = (m+1)^{\text{th}}$ prompt
    Choose parameters λ and μ using cross validation over first m prompts by solving (5) (or (6)) using CVX
    Obtain parameters $\boldsymbol{\beta}_0, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ by solving (5) (or (6)) using CVX
    Predict the ratings at prompt $p = m+1$ using (3) (or (4))

Setting $\lambda = 0$ in Eqs. (5) and (6) reduces them to ordinary network-agnostic least-squares regression models. In simple terms, a participant's 5th prompt rating is then predicted by running linear regression over previous ratings, which is used as a baseline. For $\lambda \neq 0$ chosen by cross-validation, the network-effects are switched on in Eqs. (5) and (6), and the optimization procedure determines parameters that minimize the total variation of the participant ratings in combination with least-squares regression. Peer effects are then determined by comparing the model with $\lambda \neq 0$ to that with $\lambda = 0$. Relative prediction errors defined as $\left\| \widehat{\boldsymbol{\beta}}_0 + 5\widehat{\boldsymbol{\beta}}_1 - \mathbf{r}^{(5)} \right\| / \left\| \mathbf{r}^{(5)} \right\|$ and $\left\| \widehat{\boldsymbol{\beta}}_0 + 5\widehat{\boldsymbol{\beta}}_1 + \sqrt{5}\widehat{\boldsymbol{\beta}}_2 - \mathbf{r}^{(5)} \right\| / \left\| \mathbf{r}^{(5)} \right\|$ respectively for Eqs. (5) and (6) are reported in Table 1.

Modeling and simulation. We proceed to develop a model for the participant ratings' evolution which facilitates simulation of the network process. We model the temporal update of the ratings $\mathbf{r}^{(p)}$ as a diffusion process on the graph G, where the trade of feedback can take place without decreasing the individual level of knowledge/speaking ratings. We impose a positive drift to the ratings in order to model the fact that a learner can accumulate an understanding of skills (i.e., build mental models of better practices in speaking) from various sources such as automated feedback, peer feedback, experience of watching peers' videos, and other external resources. In a real network, people will have different learning rates, their individual talents will differ, the external inflow of information will not be consistent, and in some less probable cases, they may forget information as well. To lump all these noisy yet positively skewed effects into one variable, we

Table 1 Fifth prompt prediction errors comparison between consensus-based ($\lambda \neq 0$) and network-agnostic ($\lambda = 0$) frameworks, for regression models with (a) linear features using Eq. (3) and (b) non-linear features using Eq. (4)

	(a) Prediction errors with linear features		(b) Prediction errors with non-linear features	
	Consensus-based ($\lambda \neq 0$)	Network-agnostic ($\lambda = 0$)	Consensus-based ($\lambda \neq 0$)	Network-agnostic ($\lambda = 0$)
Group 1	**9.85%**	14.52%	**7.83%**	8.18%
Group 2	**10.93%**	18.67%	**10.72%**	12.75%
Group 3	**12.42%**	15.62%	**9.95%**	13.14%
Group 4	**11.26%**	15.4%	**11.31%**	12.81%
Group 5	**12.29%**	13.99%	**10.02%**	12.57%
Group 6	**7.04%**	11.43%	**7.06%**	10.47%

Bold errors denote better prediction results. The optimization functions for the regression models are elaborated through Eqs. (5) and (6) in the Methods section

model the positive drift ϵ as a Gaussian random variable with a positive mean $\mu > 0$. Superposition of random effects is well modeled by a Gaussian random variable by virtue of the Central Limit Theorem. The ratings received by participants lie between 1–5, so we introduce an explicit control procedure to bound the ratings in our model. Thus the evolution of ratings can be modeled via the Laplacian dynamics

$$\mathbf{r}^{(p+1)} = \mathcal{P}_{[r_{\min}, r_{\max}]}\left(\mathbf{r}^{(p)} - c\mathbf{L}^{(p)}\mathbf{r}^{(p)} + \epsilon \right), \qquad (7)$$

where $\mathbf{r}^{(p)}$ is the ratings vector of pth prompt, $c \in (0, 1/d_{max})$ is the diffusion constant and d_{max} is the maximum degree of nodes at the corresponding prompt, $\mathbf{L}^{(p)}$ is the Laplacian matrix of the graph G at the pth prompt, ϵ is a Gaussian random vector with mean $\mu > 0$ and given variance σ^2, and $\mathcal{P}_{[r_{\min}, r_{\max}]}(\cdot)$ is a projection operator onto the interval $[r_{\min}, r_{\max}]$. For ROC Speak one has $r_{\min} = 1$ and $r_{\max} = 5$.

Figure 3 demonstrates the overall idea pictorially. To understand the chosen dynamics, disregard the projection operator in Eq. (7) for the sake of a simpler argument. Then notice that the update $\mathbf{r}^{(p+1)} = \mathbf{r}^{(p)} - c\mathbf{L}^{(p)}\mathbf{r}^{(p)}$ represents a Laplacian-based network diffusion process (Olfati-Saber and Murray 2004), where the future rating of a given participant depends on the ratings of his/her peers in the current prompt ($\mathbf{r}^{(p)}$) and the nature of the interactions taking place in the learning community ($\mathbf{L}^{(p)}$). Focusing on the ith participant recursion in Eq. (7) (modulo the projection operator), one obtains the scalar update

$$r_i^{(p+1)} = (1 - cd_i)r_i^{(p)} + c\sum_{j \in \mathcal{V}} W_{ij}^{(p)} r_j^{(p)} + \epsilon_i, \qquad (8)$$

where $r_i^{(p)}$ is ith participant rating at the pth prompt. In obtaining Eq. (8), we have used the definition of graph Laplacian, i.e., $\mathbf{L}^{(p)} := \mathbf{D}^{(p)} - \mathbf{W}^{(p)}$. The quantity $(1 - cd_i)r_i^{(p)} + c\sum_j W_{ij}^{(p)} r_j^{(p)}$ is a weighted average of participant i and his/her neighbors' ratings, with each neighbor's weights being proportional to the number of interactions that s/he has with participant i. This way, the model captures peer-influenced buildup of ratings across the network where the diffusion constant c is a relatively small number. Further intuition can be gained by interpreting $\mathbf{r}^{(p+1)} = \mathbf{r}^{(p)} - c\mathbf{L}^{(p)}\mathbf{r}^{(p)}$ as a gradient-descent iteration to minimize the total variation functional $\mathrm{TV}(\mathbf{r}) := \mathbf{r}^T \mathbf{L}^{(p)} \mathbf{r}$ in Eq. (1). This suggests that Eq. (7) will drive the ratings towards a consensus of minimum total variation.

We test the proposed model in Eq. (7) by running a numerical simulation that synthetically generates participant ratings across prompts. We synthetically generate graphs with structural properties resembling our dataset. Given that in our platform, participants can, in principle, send communication at random, we generate Erdös–Rényi random graphs for each prompt with the average number of nodes in the communities ($N = 26$ nodes) (Erdos and Rényi, 1960). Each edge is included in the graph with probability $p = 0.5$, so that the expected number of edges $\binom{N}{2} p$ matches the number of edges in the dataset. We set the positive drift term ϵ to have a mean $\mu = 0.05$ and standard deviation $\sigma = 0.1$ based on the ROC Speak average ratings behavior, and select $c = 0.01$.

Results

Convergence of community-wide performance with increasing interactions. We evaluate the total variation of the rating signals $\mathbf{r}^{(p)}$ across the ROC Speak network at the end of each prompt [cf. (1)]. We collect ith community's TV of the 5 prompts in a vector $\mathbf{TV}_i = \left(\text{TV}_i^{(1)}, \text{TV}_i^{(2)}, \dots, \text{TV}_i^{(5)} \right)^T$ and plot the normalized total variation as $\mathbf{TV}_i / \|\mathbf{TV}_i\|$ over prompts in Fig. 4a, where $\|\cdot\|$ maps a vector to its Euclidean norm. The normalized total variations mostly decrease from the third prompt onwards. The dashed plot in Fig. 4a indicates the linear trend of the average across the six groups $\left(\frac{1}{6} \sum_{i=1}^{6} \mathbf{TV}_i \right)$. A diminishing trend is apparent, with a moderate yet significant slope of -0.04 ($p < 0.01$), suggesting that interacting participants gradually converge in terms of ratings.

In addition, we plot the network-wide average ratings defined in Eq. (2) as a function of prompts p in Fig. 4b. It is evident that all six groups have a trend of improvement in average ratings across prompts. The dashed plot indicates the average linear trend estimation across all six groups, which has a moderate and significant slope of 0.07 ($p < 0.05$). Taken together, these findings provide strong evidence that communities gradually approach homogeneity in terms of (improved) performance with increasing interactions among participants.

Improved prediction of ratings as a function of incorporating peer effects. As explained in the Methods section, we compare the predictions against the original ground-truth ratings of the final prompt to report the prediction errors. The results are summarized in Table 1. The first and third columns of the table show the prediction errors obtained by our proposed consensus-based predictors in Eqs. (5) and (6). The second and fourth columns refer to the baseline predictors of ordinary least-squares regression, which are computed by making the smoothness regularization term $\lambda = 0$ in Eqs. (5) and (6). Table 1 shows that the consensus-based predictors ($\lambda \neq 0$) outperform their baseline counterparts ($\lambda = 0$) in all six communities. In particular, the consensus-based approach relatively improves the predictions by 23% averaged over six groups. In other words, the ratings of one's peers and the associated interaction patterns help predict his/her performance ratings better. This, in turn, supports the idea that interactions with the neighbors indeed impact one's learning outcomes.

Fig. 3 An example demonstrating how participant 1's rating in the $(p+1)$th prompt, $r_1^{(p+1)}$, is calculated from the participant's own rating in the pth prompt, $r_1^{(p)} = 4$, single interactions with participants 2, 3 and 4, and a positive drift ϵ_1. The pth prompt ratings for the participants are shown by blue bars. Here, participants 2 and 4 both have ratings of 3 in the pth prompt, and participant 3 has a rating of 5. The net diffusion effect with peers takes participant 1's rating from 4 down to 3.99 (denoted by a yellow bar). However, a Gaussian random positive drift $\epsilon_1 = 0.0676$ (denoted by a green bar on top of the yellow bar) represents a sample realization of participant 1's gathering of understanding of speaking skills from sources external to peer feedback, and pushes his rating in the $(p+1)$th prompt to $r_1^{(p+1)} = 3.99 + 0.0676 = 4.0576$. The magnified plot reflects the probability distribution of participant 1's rating at $(p+1)$th prompt which is centered around $3.99 + \mu$, with mean μ. This also shows that the rating in $(p+1)$th prompt can take any value in the interval $[r_{min}, r_{max}]$, but with a higher probability it is closer to $3.99 + \mu$
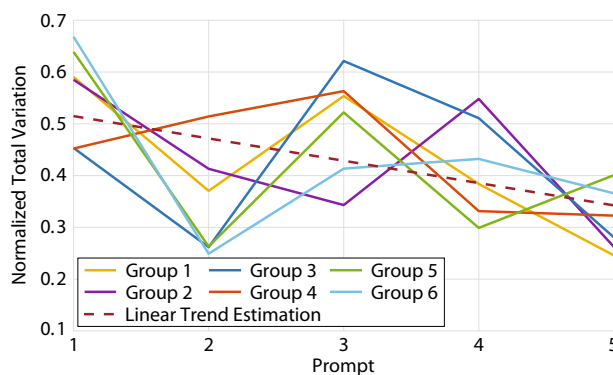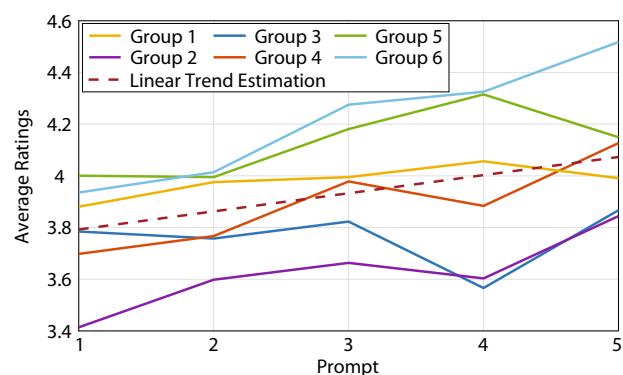

(a) Total variation of ratings across networks versus prompts.

(b) Average ratings of the participant groups versus prompts.

Fig. 4 Total variations and average ratings as functions of prompts. In **a**, the total variation decreases across prompts while **b** shows the network-wide average ratings increase. The dashed lines are the linear trend estimations across prompts

Capturing the dynamics of the interactions. To simulate the proposed Laplacian dynamics-based model described in Eq. (7), we run iterations with $\mathbf{r}^{(1)}$ as initialization, where $r_i^{(1)}$'s are drawn uniformly at random from the interval Eqs. (1, 5). Figure 5a shows two realizations of the total variation Eq. (1) as it evolves over prompts, superimposed to the mean evolution obtained after averaging 1000 independent such realizations. The specific realizations show fluctuations similar to those observed in Fig. 4a, but the trend is diminishing towards zero meaning that ratings converge with temporal evolution. Likewise, Fig. 5b shows two realizations and the mean evolution of the network-wide average ratings $\bar{r}^{(p)}$ in Eq. (2) as a function of prompts.

The figures indicate that the simulation results accurately capture the positively skewed effects of interactions in the gradual buildup of speaking skills. While the network model is admittedly simple, the synthesized sample paths qualitatively match the trends in the data. This is further illustrated in Fig. 6, where participant-ratings are shown as blue bars demonstrating a decrease in the variation of ratings as the simulation evolves in time, with the average ratings approaching the limiting value $r_{\max}$.

Discussion
In this paper, we explored how participants' performance ratings progressed as they interacted in online learning communities. In particular, we asked whether the learners' ratings were affected by their interaction patterns and their peers' ratings, and also whether the communities gradually approached homogeneity in speaking performance ratings. We found positive evidence for both of the questions.

Developing competence in speaking skills has its importance across professional and personal settings. There have been some

explorations in directional peer tutoring when it comes to developing oral expression in learning a second language—namely between two students in school or between a family member and a student at home (Duran et al., 2016; Topping et al., 2017). However, in the setting of small learning groups, the network-effects in developing one's speaking skills received negligible attention in literature—a context we presented our insights in. Our novel dataset of six online communities thus allowed us to study the peer-effects in an objective manner.

We studied performance ratings as graph signals on top of comment-based interaction networks. The emerging field of GSP has received attention from a variety of fields such as Neuroscience and medical imaging, but GSP was not applied to a learning setting previously. Taking a GSP-based prediction approach helped us observe that the learners' ratings had a connection with their interaction patterns, as well as the rating signals of the peers they interacted with. Both parts of this observation agree with previous findings in other scopes of learning. For instance, Fidalgo-Blanco et al., (2015) had shown that there is a direct relation between quantities of interaction and individual performance outcomes. Using Texas Schools Microdata, it was shown that an exogenous change of 1 point in peers' reading scores raises a student's own score between 0.15 and 0.4 points (Hoxby, 2000). Our prediction formulation outperformed the baseline relatively by 23%, thus corroborating the same insights in the context of speaking skill development.

Moreover, evidence for a gradual buildup of skills towards community-wide homogeneity was not reported in earlier learning community literature, and that is a gap we addressed. Theoretically, this effect has long been anticipated. For instance, Cowan and Jonard (2004) simulated diffusion of knowledge and suggested its gradual buildup in networks. Following the
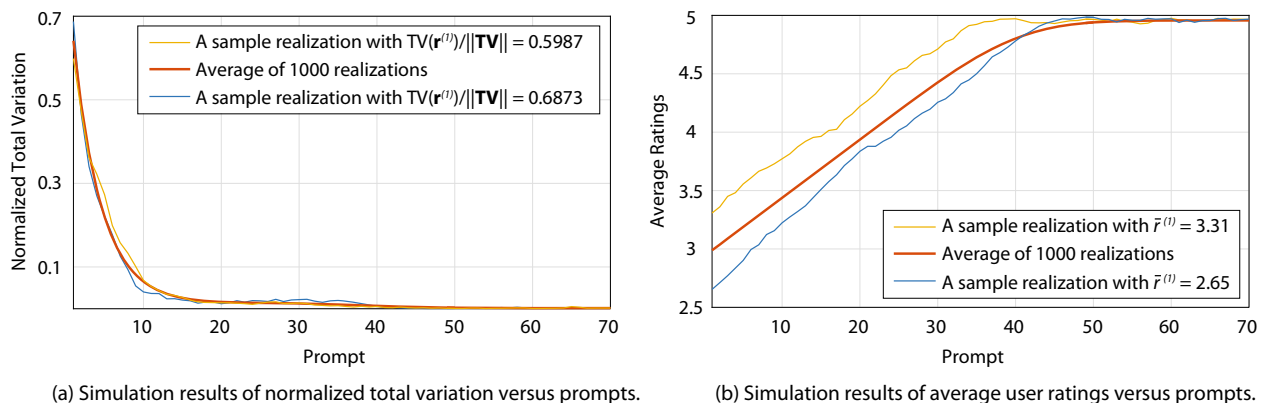

(a) Simulation results of normalized total variation versus prompts.

(b) Simulation results of average user ratings versus prompts.

Fig. 5 Simulation results of normalized total variations and average ratings as functions of prompts. The normalized total variation across the network decreases in **a**, while the average ratings saturate to the ceiling in **b**
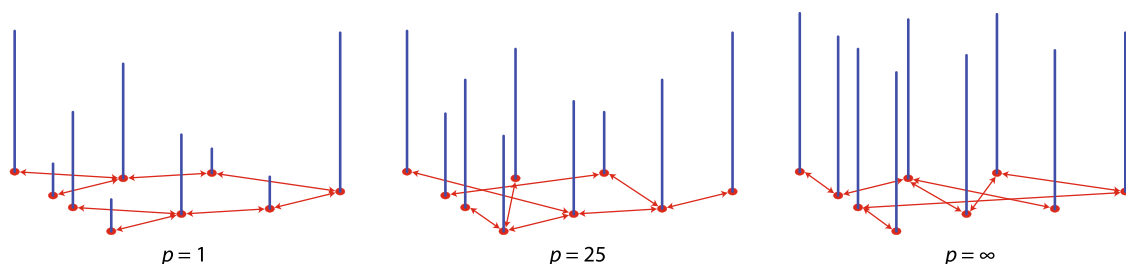

Fig. 6 A sample visualization of how the participants' ratings $\mathbf{r}^{(p)}$ (blue bars) change as total variation diminishes and average rating saturates at limiting time (i.e., prompt $p \to \infty$). The network topology changes in every prompt as the participants randomly interact with each other. A decreasing TV($\mathbf{r}$) measure is a direct manifestation of the signal $\mathbf{r}^{(p)}$ being smoother as p increases

collaborative model of learning, this buildup is also reasonable, as repetitive interactions can be anticipated to help participants develop their tacit understandings of speaking skills gradually (Davidson and Major, 2014). In our dataset, total variation of the rating signals going down (slope = −0.04, $p < 0.01$) and the average ratings going up (slope = 0.07, $p < 0.05$) with time provide evidence of skill-ratings building up among interacting participants. To the best of our knowledge, this is the first data-driven evidence for this idea. We further proposed a network diffusion model and used it to simulate the rating buildup process. The model is simple and derived from previous work (Olfati-Saber and Murray, 2004), but we modified the formulation therein to capture the peer-learning effects as observed in our analysis of the data.

It is important to stress that while our dataset is concerned with speaking skill development and therefore our insights are presented in that context, the analysis methods are not limited to speaking-skills only. Any dataset in a learning setting that comes with quantifiable learning outcome measures (e.g., grades, scores, ratings), as well as interaction measures (e.g., comments or other synchronous/asynchronous textual discussion, audio, video, or any suitable Internet-supported media based interaction) can be modeled and analyzed using our formulations.

In this paper, we did not take into account the contents of the comments, which is left as a future work. We treated each comment as an interaction unit. This still turned out to hold information regarding skill-development in the communities, as shown through our mining of the data. Similarly, the exact specifics of the speaking skills were not considered in this paper, such as the use of gestures, vocal variety, friendliness, clarity, speaking speed etc. Rather, we used an average 'overall' rating that we assumed to be an objective abstraction of speaking performance or skills. Digging deeper into the comments and connecting them with specific attributes of speaking-skills (e.g., a comment on vocal variation leading to better vocal performance) in the future might help us uncover more insights on the peer-influence characteristics.

In his study on behavior propagation, Centola (2010) showed that health behavior spreads farther and faster across clustered-lattice networks than across corresponding random networks, since the former structure facilitates social reinforcement through redundant network ties. Valente (2012) reviewed various social intervention approaches for accelerating behavior change. Exploring the effects of various network structures and intervention strategies to maximize skill development are some of our future courses of work. We thus anticipate that the current study will open up many avenues for broader exploration of peer-influenced skill development in the future.

References

Boyd S, Vandenberghe L (2004) Convex optimization, Cambridge University Press

Brinton CG, Chiang M, Jain S, Lam H, Liu Z, Wong FMF (2014) Learning about social learning in moocs: From statistical analysis to generative model. IEEE Trans Learn Technol 7(4):346–359

Brouwer J, Flache A, Jansen E, Hofman A, Steglich C (2017) Emergent achievement segregation in freshmen learning community networks. Higher Education p 1–18

Busch P (2008) Tacit knowledge in organizational learning. IGI Global

Carolan BV (2013) Social network analysis and education: theory, methods and applications. Sage Publications

Celant S (2013) The analysis of students academic achievement: the evaluation of peer effects through relational links. Qual Quant 47(2):615–631

Centola D (2010) The spread of behavior in an online social network experiment. Science 329(5996):1194–1197

Cho H, Gay G, Davidson B, Ingraffea A (2007) Social networks, communication styles, and learning performance in a cscl community. Comput Educ 49(2):309–329

Chugh R et al. (2013) Workplace dimensions: tacit knowledge sharing in universities. J Adv Manag Sci 1(1):24–28

Chung FK (1997) Spectral graph theory. CBMS regional conference series in mathematics. No. 92

Cowan R, Jonard N (2004) Network structure and the diffusion of knowledge. J Econ Dyn Control 28(8):1557–1575

Cross KP (1998) Why learning communities? why now? Campus 3(3):4–11

Dado M, Bodemer D (2017) A review of methodological applications of social network analysis in computer-supported collaborative learning. Educ Res Rev 22:159–180

Davidson N, Major CH (2014) Boundary crossings: cooperative learning, collaborative learning, and problem-based learning. J Excellence College Teaching, 25(3, 4), 7–56

De Backer L, Van Keer H, Valcke M (2015) Exploring evolutions in reciprocal peer tutoring groups' socially shared metacognitive regulation and identifying its metacognitive correlates. Learn Instr 38:63–78

Deri JA, Moura JMF (2016) New york city taxi analysis with graph signal processing. 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP). pp 1275–1279

Duran D, Flores M, Oller M, Thomson-Garay L, Vera I (2016) Reading in pairs. Peer tutoring for reading and speaking in English as a foreign language. http://grupsderecerca.uab.cat/grai/en/node/3882

Erdos P, Rényi A (1960) On the evolution of random graphs. Publ Math Inst Hung Acad Sci 5(1):17–60

Fidalgo-Blanco Á, Sein-Echaluce ML, García-Peñalvo FJ, Conde MÁ (2015) Using learning analytics to improve teamwork assessment. Comput Human Behav 47:149–156

Frank KA, Muller C, Schiller KS, Riegle-Crumb C, Mueller AS, Crosnoe R, Pearson J (2008) The social dynamics of mathematics coursetaking in high school. Am J Sociol 113(6):1645–1696

Friedman J, Hastie T, Tibshirani R (2001) The elements of statistical learning, Vol. 1, Springer Series in Statistics. Springer, Berlin

Fung M, Jin Y, Zhao R, Hoque ME (2015) Roc speak: Semi-automated personalized feedback on nonverbal behavior from recorded videos. Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, pp 1167–1178

Gale D, Kariv S (2003) Bayesian learning in social networks. Games Econ Behav 45(2):329–346

Gašević D, Zouaq A, Janzen R (2013) "Choose your classmates, your gpa is at stake!" the association of cross-class social ties and academic performance. Am Behav Sci 57(10):1460–1479

Golub B, Sadler ED (2017) Learning in social networks

Grant M, Boyd S, Ye Y (2008) Cvx: Matlab software for disciplined convex programming

Harasim L (2017) Learning theory and online technologies. Taylor & Francis

Hoxby C (2000) Peer effects in the classroom: Learning from gender and race variation. Technical report. National Bureau of Economic Research

Huang W, Goldsberry L, Wymbs NF, Grafton ST, Bassett DS, Ribeiro A (2016) Graph frequency analysis of brain signals. IEEE J Sel Top Signal Process 10(7):1189–1203

Huitsing G, Veenstra R, Sainio M, Salmivalli C (2012) "It must be me" or "it could be them?": the impact of the social network position of bullies and victims on victims' adjustment. Soc Netw 34(4):379–386

Johnson DW, Johnson RT (2009) An educational psychology success story: social interdependence theory and cooperative learning. Educ Res 38(5):365–379

Jones C (2015) Networked learning: an educational paradigm for the age of digital networks. Springer

Kotzagiannidis MS, Dragotti PL (2016) The graph fri framework-spline wavelet theory and sampling on circulant graphs. Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on, IEEE. pp 6375–6379

Leidner DE, Jarvenpaa SL (1995) The use of information technology to enhance management school education: a theoretical view. MIS Quarterly. pp 265–291

Lomi A, Snijders TA, Steglich CE, Torló VJ (2011) Why are some more peer than others? evidence from a longitudinal study of social networks and individual academic performance. Soc Sci Res 40(6):1506–1520

Mayes T, De Freitas S (2013) Technology-enhanced learning. Beetham, H. & Sharpe. pp 17–30

Newman M (2010) Networks: an introduction. Oxford University Press

Olfati-Saber R, Fax JA, Murray RM (2007) Consensus and cooperation in networked multi-agent systems. Proc IEEE 95:215–233

Olfati-Saber R, Murray RM (2004) Consensus problems in networks of agents with switching topology and time-delays. IEEE Trans Autom Control 49 (9):1520–1533

O'Malley C (2012) Computer Supported Collaborative Learning, Vol. 128. Springer Science & Business Media

Palonen T, Hakkarainen K (2013) Patterns of interaction in computersupported learning: a social network analysis. Fourth International Conference of the Learning Sciences, pp 334–339

Pang J, Cheung G (2016) Graph laplacian regularization for inverse imaging: analysis in the continuous domain. Preprint at arXiv:1604.07948

Putnik G, Costa E, Alves C, Castro H, Varela L, Shah V (2016) Analysing the correlation between social network analysis measures and performance of students in social network-based engineering education. Int J Technol Des Educ 26(3):413–437

Roscoe RD, Chi MT (2008) Tutor learning: The role of explaining and responding to questions. Instr Sci 36(4):321–350

Rui L, Nejati H, Safavi SH, Cheung N-M (2017) Simultaneous low-rank component and graph estimation for high-dimensional graph signals: application to brain imaging. Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on, IEEE. pp 4134–4138

Russo TC, Koesten J (2005) Prestige, centrality, and learning: a social network analysis of an online class. Commun Educ 54(3):254–261

Shuman DI, Narang SK, Frossard P, Ortega A, Vandergheynst P (2013) The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process Mag 30(3):83–98

Smith BL, MacGregor J, Matthews R, Gabelnick F (2009) Learning communities: reforming undergraduate education

Smith RA (2015) Magnets and seekers: a network perspective on academic integration inside two residential communities. J High Educ 86(6):893–922

Tawfik AA, Reeves TD, Stich AE, Gill A, Hong C, McDade J, Pillutla VS, Zhou X, Giabbanelli PJ (2017) The nature and level of learner-learner interaction in a chemistry massive open online course (mooc). J Comput High Educ 29 (3):411–431

Thanou D, Chou PA, Frossard P (2016) Graph-based compression of dynamic 3d point cloud sequences. IEEE Trans Image Process 25(4):1765–1778

Topping K, Buchs C, Duran D, Van Keer H (2017) Effective peer learning: from principles to practical implementation, Taylor & Francis

Valente TW (2012) Network interventions. Science 337(6090):49–53

Data availability

The dataset and the analysis codes can be found in: https://github.com/ROC-HCI/ROCSpeak-Buildup-of-Skills

Author contributions

R.S., R.A.B., M.K.H., G.G., G.M., and M.E.H. designed and performed the research and wrote the manuscript. All the figures are prepared by the authors.

Reprints and permission information is available online at http://www.nature.com/reprints

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims inpublished maps and institutional affiliations.

Additional information

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://www.nature.com/reprints

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims inpublished maps and institutional affiliations.