Predicting evaluations of entrepreneurial pitches based on multimodal nonverbal behavioral cues and self-reported characteristics

Kostas Stoitsas
k.stoitsas@tilburguniversity.edu
Department of Cognitive Science and Artificial Intelligence, Tilburg University
Tilburg, North Brabant, The Netherlands

Itır Önal Ertuğrul
Department of Information and Computing Sciences, Utrecht University
Utrecht, The Netherlands
i.onalertugrul@uu.nl

Werner Liebregts
w.j.liebregts@tilburguniversity.edu
Jheronimus Academy of Data Science (JADS)
‘s-Hertogenbosch, North Brabant, The Netherlands

Merel M. Jung
Department of Cognitive Science and Artificial Intelligence, Tilburg University
Tilburg, The Netherlands
m.m.jung@tilburguniversity.edu

ABSTRACT

Acquiring funding for a startup venture often involves pitching a business idea to potential investors. Insight into the nonverbal behavioral cues that impact the investment decision making process can help entrepreneurs to improve their persuasion skills and can provide valuable insights to investors and researchers. Previous research on the prediction of investment decisions in entrepreneurial pitches has primarily focused on analyzing (usually unimodal) behavioral cues from pitchers only. To address this gap, in this study we compare the predictive performance of different feature sets consisting of nonverbal behavior cues from different modalities (i.e., facial expressions, head movement, and vocal expressions) from both pitchers and investors and their self-reported characteristics. Our findings show promising results for the prediction of investor's evaluations of entrepreneurial pitches. Multimodal behavioral cues, especially head movement and vocal expressions, were found to be most predictive.

CCS CONCEPTS

• **Computing methodologies → Biometrics**.

KEYWORDS

Entrepreneurial pitch competition, Decision making process, Social signal processing, Nonverbal behavior, Multimodal interaction

ACM Reference Format:
Kostas Stoitsas, Itır Önal Ertuğrul, Werner Liebregts, and Merel M. Jung. 2022. Predicting evaluations of entrepreneurial pitches based on multimodal nonverbal behavioral cues and self-reported characteristics. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '22 Companion), November 7–11, 2022, Bengaluru, India.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3536220.3558041

1 INTRODUCTION

Nowadays for entrepreneurs the process of getting funding commonly involves a pitch of their business idea to potential investors. The success of the pitch does not depend exclusively on the content of the presentation, personal characteristics such as demographics, entrepreneurial traits, and competencies might also play a significant role in the outcome [9, 11]. Additionally, the nonverbal behavior of pitchers such as facial and vocal expressions, gestures, body posture, eye contact, and gaze might be potential predictors of their success [12, 17, 19].

In the crowdfunding domain, several studies have explored predictors of successfully funded projects. For example, [29] presented promising results for the prediction of success of crowdfunding projects based on a combination of different feature sets including project and founder characteristics, text descriptions, and audio features extracted from promotion videos by leveraging a pre-trained VGG-ish deep neural network. In a study on facial expressions displayed in pictures on crowdfunding pages, results showed that moderate display of expressions of happiness and sadness had a positive effect on funding decisions [27]. In contrast, showing high intensity emotions was found to have a negative impact.

In the context of entrepreneurial pitches, a similar inverted U-shaped relation was found between the expression of joy by pitchers and the attraction of funding [19]. Another study expanded on these findings, showing inverted U-shaped patterns for frequencies of facial expressions of happiness, anger, and fear whereas expressions of sadness were found to have an negative impact on funding decisions [35]. Moreover, the use of gestures to convey a business idea was found to have a positive effect on persuading investors while the use of literal or figurative language to describe the business idea had little impact [12]. These findings suggest that behavioral cues, when used appropriately, can be used to grasp the attention of the investors.

ICMI '22 Companion, November 7–11, 2022, Bengaluru, India

ACM ISBN 978-1-4503-9389-8/22/11…$15.00
https://doi.org/10.1145/3536220.3558041

Previous research on the prediction of investment decisions in entrepreneurial pitches has primarily focused on analyzing (usually unimodal) behavioral cues from pitchers only [17]. In contrast, in studies on predicting hiring decisions based on employment interviews, multimodal behavior expressed by both parties have been considered. For example, multimodal behavioral cues expressed by both the job applicant and interviewer during a job interview have shown to be predictive of hirability [25]. Audio cues from applicants regarding speaking activity (e.g., use of short utterances) and prosody and visual cues from the interviewer (e.g., head nodding) have been found to be most predictive. Interestingly, applicant's answers on questionnaires commonly used for personnel selection were found to have no predictive value. In another study on predicting hiring decision based on asynchronous video interviews, multimodal models trained on verbal content, speech prosody, and facial expressions were found to outperform unimodal models [18]. These studies highlight the importance of multimodal behavioral cues of both parties for understanding decision making processes based on presentations.

To address this gap the contributions of the current study are two-fold: 1) in our prediction models we also include extracted behavioral cues of investors and from interactions between pitcher-investor pairs; 2) cues from different modalities are included. The predictive performance of different feature sets are compared, consisting of nonverbal behavior cues from different modalities (i.e., facial expressions, head movement, and vocal expressions) from both pitchers and investors and their self-reported characteristics. Models for the automatic prediction of invest decisions can be integrated into training applications to help entrepreneurs to improve their pitching skills and can provide investors and researchers with insights into the decision making process. Our findings show promising results for the prediction of investor's evaluations of entrepreneurial pitches. Multimodal behavioral cues, especially head movement and vocal expressions, were found to be most predictive whereas self-reported characteristics were found to be least predictive.

2 METHOD

2.1 Entrepreneurial pitch data-set

In this study we used a data-set containing video recordings from entrepreneurial pitch competitions and survey data from the pitchers and investor panels [22]. Data was collected with the aim of improving our understanding on the decision making process in an entrepreneurial context by applying modern-day data science techniques [21]. The data collection and management process has been approved by the ethics committee of the Tilburg School of Economics and Management.

2.1.1 Video data. The data-set included video recordings from individuals who perform an entrepreneurial pitch on behalf of their team about their start-up business idea and from a panel of up to three investors who are assessing the pitches. The pitchers were university students who took part in the pitch competition as part of a course on entrepreneurship in data science. The investors were all professionals with extensive experience in the industry. Pitchers had a maximum of three minutes to perform their pitch. An interactive session with a maximum of ten minutes followed directly after the pitch where the investors could ask questions to the pitcher. For the purpose of this study, only the video recordings of the three minute pitch sessions were considered. In total, we included 20 pitches from three pitch sessions involving a total of 53 pitcher-investor interactions. Pitch session 1 contained 7 pitches and a panel of 3 investors; pitch session 2 included 7 pitches and 2 investors; and pitch session 3 included 6 pitches and 3 investors.

2.1.2 Survey data. The data-set also contained self-reported survey data as various characteristics (e.g., personality, experiences, and attitudes) have been found to affect investment decisions [9, 11]. An overview of the characteristics of the pitchers and investors that were included as predictors for evaluating pitch outcomes is provided in Table 1. The HEXACO Personality Inventory-Revised was used to assess six major personality dimensions for both pitchers and investors: honesty-humility, emotionality, extroversion, agreeableness (versus anger), conscientiousness, openness to experiences [1]. Moreover, questions were included to examine characteristics of pitchers and investors related to empathy: fantasy, empathic concern, and personal distress [13]. Additionally, pitchers answered questions regarding their general self-efficacy [28], specific self-efficacy [23] (searching, planning, marshaling, people aspects of implementation, financial aspects of implementation), entrepreneurial passion [7], coach-ability [10], and characteristics of their team (heterogeneity [31], flexibility [6], satisfaction [32], efficacy [15], viability [4]). Investors answered questions about their professional experience including the number of (co-)founded ventures and the total years of involvement in these ventures. Moreover, the attitude towards risk [14] and the imaginativenes [20] (creative imagination, social imagination, practical imagination) of investors was evaluated. In addition, we included the investor's age and years of university-level education. Pitchers completed all questionnaires before the day of the pitch whereas investors completed their questionnaires after the pitches. Ratings on multi-item scales were averaged as they were found to be internally consistent (Cronbach's $\alpha \geq .72$). Negatively-phrased questions were reverse scored before calculating averages.

2.2 Multimodal behavioral feature extraction

Besides the self-reported characteristics listed in Table 1, multimodal behavioral features were extracted from the video recordings (see Table 2 for an overview). These features included behaviors expressed by both the person pitching the startup idea and the investors who are listening to the pitch. In addition, features regarding the interaction dynamics were included to capture synchrony or mimicry behavior between pitcher-investor pairs which could indicate social affiliation or act as a mechanism to facilitate the contagiousness of entrepreneurial passion [21].

2.2.1 Facial expressions. Facial action units (AUs) were extracted from the video recordings of the pitchers and investors using OpenFace 2.0 [2]. The activation of different AUs were combined to infer the intensity of different emotional expressions: for happiness: AU6 (Cheek Raiser) and AU12 (Lip Corner Puller); for sadness: AU1

Table 1: Self-reported characteristics of pitchers and investors used as predictors for pitch evaluations.

Pitcher's characteristics		Investor's characteristics	
Honesty-Humility	Self-efficacy: planning	Honesty-Humility	No. of years funding a new venture
Emotionality	Self-efficacy: marshaling	Emotionality	Years at university
Extroversion	Self-efficacy: people implementation	Extroversion	Age
Agreeableness	Self-efficacy: financial implementation	Agreeableness	Entrepreneurial experience
Conscientiousness	Entrepreneurial passion	Conscientiousness	Attitude towards risk
Openness to experiences	Coach-ability	Openness to experiences	Creative imagination
Empathy-fantasy	Team heterogeneity	Empathy-fantasy	Social imagination
Empathic concern	Team flexibility	Empathic concern	Practical imagination
Personal distress	Team satisfaction	Personal distress	
General self-efficacy	Team efficacy	No. of funded ventures	
Self-efficacy: searching	Team viability	No. of years running a new venture	

Table 2: Overview of the extracted behavioral features and characteristics.

Modality	Feature description	Extract from ...
Facial expressions	Recurrence rate of expression of happiness, sadness, surprise	Pitcher-investor pairs
	Percentage of smiling and expression of surprise	Investors
	Percentage of smiling	Pitchers
Head movement	Recurrence rate of head movements in x, y, and z direction	Pitcher-investor pairs
	Time series features of movement in x and y direction	Investors
	Time series features of movement in x and y direction	Pitchers
Vocal expressions	Low-level descriptors and functionals (OpenSmile)	Pitchers
	Deep context-aware features (VGG-100)	Pitchers
Characteristics	Self-reported characteristics	Pitchers
	Self-reported characteristics	Investors

(inner brow raiser), AU4 (brow lowerer), and AU15 (lip corner depressor); and for surprise: AU1, AU2 (outer brow raiser), AU5 (upper lid raiser), and AU26 (jaw drop). Facial expression features included recurrence rates of expressions of happiness, sadness, and surprise for all pitcher-investor pairs. Recurrence rates were calculated per pitcher-investor pair using the CRQA R package [36]. In general, recurrent rates between 0 and 50% were found, possibly indicating synchrony and mimicry behavior between pitchers and investors. Moreover, the occurrence of smiling by both pitchers and investors and investor's expressions of surprise were calculated based on the simultaneous activation of the corresponding AUs: AU6 and AU12 for smiling and AU1, AU2, AU5, and AU26 for surprise. Percentages were used for normalization to account for differences in pitch duration.

2.2.2 Head movement. Time series of head positions in the x and y axes as detected by OpenFace [2] were fed into the TSFEL Python library [3] to extract head movement features. In total, TSFEL extracts over 60 different features per time series in the statistical, temporal, and spectral domains. Additionally, recurrent rates of head movement of all pitcher-investor pairs were calculated based on OpenFace's three dimensional head position data. Multidimensional cross-recurrence quantification analysis was performed with optimization of the radius, embedding dimension, and delay [34].

2.2.3 Vocal expressions. Both deep audio features and hand-crafted audio features were extracted from the videos of the pitchers as previous work has shown that a combination of these features were most predictive of investment decisions [17]. Videos of the pitchers were converted from mp4 to wav files by using the MoviePy Python library and then converted to spectrograms using the SciPy library [33]. Deep context-aware features were extracted from the spectogram images (size 395 x 574 x 3) using the VGG16 architecture pretrained with ImageNet [30]. The VGG model, which is a convolutional neural network-based architecture, has proven to be effective for image classification tasks. The output of the VGG model was set to extract 100 features. In addition to the deep features, the OpenSmile library was used to extract 62 features from GeMAPSv01b to capture Low-Level Descriptors (LLD) and functionals [16].

2.3 Predicting pitch evaluations

2.3.1 Responses for prediction. After each pitch, the investors completed a survey concerning the quality of the entrepreneurial idea. More specifically, investors were asked to evaluate the pitch idea with respect to the "probability that you would invest in the idea", "overall quality of business idea", and "likelihood that the business idea could serve as the basis of a successful new venture". The range of these responses was between 0 and 100.

Pearson correlation coefficients showed high positive correlations between all three responses: "probability that you would invest in the idea" and "overall quality of business idea" ($r(51) = .89, p < .001$); "probability that you would invest in the idea" and "likelihood that the business idea could serve as the basis of a successful new venture" ($r(51) = .84, p < .001$); and "overall quality of business idea" and "likelihood that the business idea could serve as the basis of a successful new venture" ($r(51) = .81, p < .001$). These strong correlations are expected as all three responses measure the evaluation of the pitch, albeit at different levels. As a result, different factors might have influenced the three pitch evaluations and therefore all three responses were included as separate outcome predictors in this study. To illustrate, it is expected that investors' probability of investment is informed by their own judgment whether the business idea could serve as the basis of a successful new venture. However, the business idea must also fit the investor's portfolio. Similarly, the quality of the business idea can affect both the probability of investment and the estimated likelihood that the business idea could turn into a successful new venture. Nevertheless, no matter how high the (perceived) overall quality of a business idea is, it could still be considered an idea with only little promise to become a viable business. This also strongly depends on the perceived capabilities of the founding team members and the targeted market.

2.3.2 Model training and evaluation. XGBoost (XGB) regression models [8] were trained for the prediction of the three pitch evaluation responses. Since we deal with few data samples and many features, nested 3-fold cross-validation was performed with grid search for hyperparameter optimization [26]. Each of the three pitch sessions was alternately used as test set while the other two were used for training and optimizing the XGB models. Pitch session 1 included 21 interactions (7 pitchers × 3 investors), pitch session 2 included 14 interactions (7 pitchers × 2 investors, and pitch session 3 included 18 interactions (6 pitchers × 3 investors). XGB models were optimized by tuning the learning rate (0.05-0.9), loss function (squared error, absolute error, Huber), the criterion to measure the quality of a split (friedman mse, squared error, mean squared error), and the minimum samples per leaf (1-50). Optimization of XGB occurs by minimizing the Mean Absolute Error (MAE). In addition, Root Mean Squared Error (RMSE) was reported to study the effect of outliers. Multimodal late fusion predictions were obtained by averaging the decisions of the regressors. The average MAE and RMSE were calculated based on all combined multimodal behavioral features and on all combined multimodal behavioral features plus the self-reported characteristics.

3 RESULTS

Performance of the models for predicting the three pitch evaluation responses based on different feature sets was evaluated and compared. The results for predicting the probability of investment are presented in Table 3. The best performance was achieved for the models trained on head movement and vocal expression features. In Table 4 the results for predicting the overall quality of business idea are presented. Again, the models based on head movement and vocal expression features provided the best results. Finally, regression models for predicting the likelihood that the business idea could serve as the basis of a successful new venture showed the best performance for the model trained on the vocal expression features followed by the multimodal late fusion modal that combined all behavioral features (see Table 5).

In general, average MAE values were lower than the average RMSE. However the average values were in the same order indicating the absence of outliers in our data. The overall best performance was achieved for the model predicting the overall quality of the business idea using the head movement features. Moreover, the results showed that models trained on the self-reported characteristics performed worse than models trained on the different behavioral feature sets. As a result, multimodal late fusion models including all behavioral features performed better than models including all behavioral features plus the self-reported characteristics.

4 DISCUSSION

The goal of the current study was to investigate the importance of behavior cues from different modalities (i.e., facial expressions, head movement, and vocal expressions) expressed by pitchers and investors and their self-reported characteristics for predicting funding decisions. Comparable performances were obtained for the three different responses related to the evaluation of the pitched business idea. For all cases, head movement and vocal expressions were found to be most predictive of pitch evaluations whereas self-reported characteristics were found to be least predictive.

Our results are in agreement with previous work, showing that head nodding from the audience can improve the confidence of a speaker and as a result can increase their persuasiveness [5]. Similarly, vocal expressions have been found to play an important role in the persuasiveness of the speaker and for conveying their message [24]. Surprisingly, the characteristics of both the pitcher (e.g., extroversion and entrepreneurial passion) and the investor (e.g., empathy and attitude towards risk) were found to not improve the predictive capability of the models [21]. A possible reason could be that the self-reported characteristics were biased, reflecting a more idealized version of the person.

Future research will focus on an in-depth evaluation of the most important features within the feature sets to get more insight in the contribution of behavioral cues from pitchers and investors as well as their interaction dynamics. Moreover, previous research has shown that the stage of multimodal fusion can impact model performance [18], therefore a future step will be the application of early fusion strategies.

A limitation of our study is the fact that our models were trained on a small dataset. An extended dataset with more pitches might improve prediction performance. In addition, pitches were performed by students as part of a course on entrepreneurial startups. As a consequence, the generalizability of the model to pitches performed by professional entrepreneurs will need to be evaluated.

In conclusion, the current study provides promising results for the prediction of pitch evaluations based on behavioral cues from pitchers, investors, and their interaction dynamics. In the future, such prediction models can be used either by investors as supplementary tools to evaluate business ideas or by entrepreneurs to improve their persuasion skills.

Table 3: Regression results for the response "probability that you would invest in the idea" on different feature sets. MAE_{1-3} **indicate the results for the model evaluated on hold-out pitch session 1, 2, or 3, respectively.**

Modality	MAE_1	MAE_2	MAE_3	Average MAE	Average RMSE
Facial expressions	20.56	14.27	20.13	18.32	24.25
Head movement	19.13	14.28	16.01	**16.47**	**20.37**
Vocal expressions	18.03	14.23	18.69	16.98	21.29
Self-reported characteristics	20.57	14.29	21.60	18.82	24.04
Multimodal late fusion: facial expressions, head movement, vocal expressions	19.24	14.26	18.27	17.25	21.97
Multimodal late fusion: facial expressions, head movement, vocal expressions, self-reported characteristics	19.57	14.27	19.11	17.64	22.48

Table 4: Regression results for the response "overall quality of business idea" on different feature sets. MAE_{1-3} **indicate the results for the model evaluated on hold-out pitch session 1, 2, or 3, respectively.**

Modality	MAE_1	MAE_2	MAE_3	Average MAE	Average RMSE
Facial expressions	17.20	17.05	20.19	18.15	23.84
Head movement	14.35	13.53	18.22	**15.36**	**20.14**
Vocal expressions	15.14	16.88	16.62	16.21	21.18
Self-reported characteristics	19.55	18.21	18.61	18.79	25.14
Multimodal late fusion: facial expressions, head movement, vocal expressions	15.56	15.82	18.34	16.57	21.72
Multimodal late fusion: facial expressions, head movement, vocal expressions, self-reported characteristics	16.56	16.41	18.41	17.13	22.57

Table 5: Regression results for the response "likelihood that the business idea could serve as the basis of a successful new venture" on different feature sets. MAE_{1-3} **indicate the results for the model evaluated on hold-out pitch session 1, 2, or 3, respectively.**

Modality	MAE_1	MAE_2	MAE_3	Average MAE	Average RMSE
Facial expressions	18.39	17.60	19.40	18.46	23.80
Head movement	14.90	17.60	22.27	18.25	22.49
Vocal expressions	15.50	17.55	18.24	**17.09**	**21.18**
Self-reported characteristics	18.74	17.58	20.14	18.82	22.90
Multimodal late fusion: facial expressions, head movement, vocal expressions	16.26	17.58	19.97	17.93	22.49
Multimodal late fusion: facial expressions, head movement, vocal expressions, self-reported characteristics	16.88	17.58	20.01	18.15	22.59

REFERENCES

[1] Michael C. Ashton and Kibeom Lee. 2009. The HEXACO–60: A Short Measure of the Major Dimensions of Personality. *Journal of Personality Assessment* 91, 4 (2009), 340–345. https://doi.org/10.1080/00223890902935878 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/00223890902935878 PMID: 20017063.

[2] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. OpenFace 2.0: Facial Behavior Analysis Toolkit. In *2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018)*. 59–66. https://doi.org/10.1109/FG.2018.00019

[3] Marília Barandas, Duarte Folgado, Letícia Fernandes, Sara Santos, Mariana Abreu, Patrícia Bota, Hui Liu, Tanja Schultz, and Hugo Gamboa. 2020. TSFEL: Time Series Feature Extraction Library. *SoftwareX* 11 (2020), 100456.

[4] Murray R Barrick, Greg L Stewart, Mitchell J Neubert, and Michael K Mount. 1998. Relating member ability and personality to work-team processes and team effectiveness. *Journal of applied psychology* 83, 3 (1998), 377.

[5] Pablo Briñol and Richard E Petty. 2008. Embodied persuasion: Fundamental processes by which bodily responses can impact attitudes. *Embodiment grounding: Social, cognitive, affective, and neuroscientific approaches* (2008), 184–207.

[6] Michael A Campion, Ellen M Papper, and Gina J Medsker. 1996. Relations between work team characteristics and effectiveness: A replication and extension. *Personnel psychology* 49, 2 (1996), 429–452.

[7] Melissa Cardon, Denis Grégoire, Christopher Stevens, and Pankaj Patel. 2013. Measuring Entrepreneurial Passion: Conceptual Foundations and Scale Validation. *Journal of Business Venturing* 28 (05 2013), 373–396. https://doi.org/10.1016/j.jbusvent.2012.03.003

[8] Tianqi Chen and Carlos Guestrin. 2016. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA) *(KDD '16)*. ACM, New York, NY, USA, 785–794. https://doi.org/10.1145/2939672.2939785

[9] Xiao-Ping Chen, Xin Yao, and Suresh Kotha. 2009. Entrepreneur passion and preparedness in business plan presentations: a persuasion analysis of venture capitalists' funding decisions. *Academy of Management journal* 52, 1 (2009), 199–214.

[10] Michael Ciuchta, Chaim Letwin, Regan Stevenson, and Sean McMahon. 2017. Betting on the coachable entrepreneur: Signaling and social exchange in entrepreneurial pitches. *Entrepreneurship Theory and Practice* 2017 (12 2017), 15065–15065. https://doi.org/10.1177/1042258717725520_

[11] Colin Clark. 2008. The impact of entrepreneurs' oral 'pitch' presentation skills on business angels' initial screening investment decisions. *Venture Capital* 10, 3 (2008), 257–279. https://doi.org/10.1080/13691060802151945 arXiv:https://doi.org/10.1080/13691060802151945

[12] Jean S. Clarke, Joep Cornelissen, and Mark P. Healey. 2019. Actions speak louder than words: How figurative language and gesturing in entrepreneurial pitches influences investment judgments. *Academy of Management Journal* 62, 2 (Jan. 2019), 335–360. https://doi.org/10.5465/amj.2016.1008

[13] Mark H Davis et al. 1980. A multidimensional approach to individual differences in empathy. (1980).

[14] Thomas Dohmen, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G Wagner. 2011. Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the european economic association* 9, 3 (2011), 522–550.

[15] Amy Edmondson. 1999. Psychological safety and learning behavior in work teams. *Administrative science quarterly* 44, 2 (1999), 350–383.

[16] Florian Eyben, Martin Wöllmer, and Björn Schuller. 2010. Opensmile: The Munich Versatile and Fast Open-Source Audio Feature Extractor. In *Proceedings of the 18th ACM International Conference on Multimedia* (Firenze, Italy) *(MM '10)*. Association for Computing Machinery, New York, NY, USA, 1459–1462. https://doi.org/10.1145/1873951.1874246

[17] Ilona Goossens, Merel M. Jung, Werner Liebregts, and Itır Önal Ertuğrul. 2022. To invest or not to invest: Using vocal behavior to predict decisions of investors in an entrepreneurial context. In *in Proceedings of the International Workshop on Human Behavior Understanding*.

[18] Leo Hemamou, Arthur Guillon, Jean-Claude Martin, and Chloe Clavel. 2021. Multimodal Hierarchical Attention Neural Network: Looking for Candidates Behaviour which Impact Recruiter's Decision. *IEEE Transactions on Affective Computing* (2021).

[19] Lin Jiang, Dezhi Yin, and Dong Liu. 2019. Can joy buy you money? The impact of the strength, duration, and phases of an entrepreneur's peak displayed joy on funding performance. *Academy of Management Journal* 62, 6 (2019), 1848–1871.

[20] Alexander Kier and Jeffery McMullen. 2018. Entrepreneurial Imaginativeness in New Venture Ideation. *Academy of Management Journal* 61 (03 2018), amj.2017.0395. https://doi.org/10.5465/amj.2017.0395

[21] Werner Liebregts, Pourya Darnihamedani, Eric Postma, and Martin Atzmueller. 2020. The promise of social signal processing for research on decision-making in entrepreneurial contexts. *Small Business Economics: An International Journal* 55, 3 (Oct. 2020), 589–605. https://doi.org/10.1007/s11187-019-00205-1

[22] Werner Liebregts, Diemo Urbig, and Merel M. Jung. 2018-2021. Survey and video data regarding entrepreneurial pitches and investment decisions. *Unpublished raw data* (2018-2021).

[23] Jeffrey Mcgee, Mark Peterson, Stephen Mueller, and Jennifer Sequeira. 2009. Entrepreneurial Self-Efficacy: Refining the Measure. *Entrepreneurship Theory and Practice* 33 (07 2009), 965 – 988. https://doi.org/10.1111/j.1540-6520.2009.00304.x

[24] John Montrey. 2005. An Investigation Of The Effects Of Speakers' Vocal Characteristics On Ratings Of Confidence And Persuasion. *Electronic Theses and Dissertations, 2004-2019* 595 (2005). https://stars.library.ucf.edu/etd/595

[25] Laurent Son Nguyen, Denise Frauendorfer, Marianne Schmid Mast, and Daniel Gatica-Perez. 2014. Hire me: Computational inference of hirability in employment interviews based on nonverbal behavior. *IEEE transactions on multimedia* 16, 4 (2014), 1018–1031.

[26] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12 (2011), 2825–2830.

[27] Maximilian Raab, Sebastian Schlauderer, Sven Overhage, and Thomas Friedrich. 2020. More than a feeling: Investigating the contagious effect of facial emotional expressions on investment decisions in reward-based crowdfunding. *Decision Support Systems* 135 (2020), 113326.

[28] Leon Schjoedt and Justin B Craig. 2017. Development and validation of a unidimensional domain-specific entrepreneurial self-efficacy scale. *International Journal of Entrepreneurial Behavior & Research* (2017).

[29] Jiatong Shi, Kunlin Yang, Wei Xu, and Mingming Wang. 2021. Leveraging deep learning with audio analytics to predict the success of crowdfunding projects. *The Journal of Supercomputing* 77, 7 (2021), 7833–7853.

[30] Karen Simonyan and Andrew Zisserman. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition. https://doi.org/10.48550/ARXIV.1409.1556

[31] Greg L Stewart and Murray R Barrick. 2000. Team structure and performance: Assessing the mediating role of intrateam process and the moderating role of task type. *Academy of management Journal* 43, 2 (2000), 135–148.

[32] Amanuel G Tekleab, Narda R Quigley, and Paul E Tesluk. 2009. A longitudinal study of team conflict, conflict management, cohesion, and team effectiveness. *Group & Organization Management* 34, 2 (2009), 170–205.

[33] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. 2020. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* 17 (2020), 261–272. https://doi.org/10.1038/s41592-019-0686-2

[34] Sebastian Wallot. 2019. Multidimensional Cross-Recurrence Quantification Analysis (MdCRQA)–a method for quantifying correlation between multivariate time-series. *Multivariate behavioral research* 54, 2 (2019), 173–191.

[35] Benjamin J Warnick, Blakley C Davis, Thomas H Allison, and Aaron H Anglin. 2021. Express yourself: Facial expression of happiness, anger, fear, and sadness in funding pitches. *Journal of Business Venturing* 36, 4 (2021), 106109.

[36] Charles Webber and Joseph Zbilut. 2005. Recurrence quantification analysis of nonlinear dynamical systems. *Tutorials in Contemporary Nonlinear Methods for the Behavioral Sciences* (01 2005).