

This is a repository copy of *The role of face and voice cues in predicting the outcome of student representative elections*.

White Rose Research Online URL for this paper:
<https://eprints.whiterose.ac.uk/150865/>

Version: Accepted Version

Article:

Mileva, Mila orcid.org/0000-0003-0537-9702, Tompkinson, James Andrew, Watt, Dominic orcid.org/0000-0003-0978-9282 et al. (1 more author) (2020) The role of face and voice cues in predicting the outcome of student representative elections. *Personality and Social Psychology Bulletin*. pp. 617-625. ISSN 0146-1672

<https://doi.org/10.1177/0146167219867965>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

The role of face and voice cues in predicting the outcome of student representative elections

Mila Mileva¹, James Tompkinson², Dominic Watt², & A. Mike Burton¹

¹ Department of Psychology, University of York, UK

² Department of Language and Linguistic Science, University of York, UK

Correspondence to:

Mila Mileva

Department of Psychology

University of York

York

YO10 5DD, UK

mila.mileva@york.ac.uk

Running head: First impressions from faces and voices

Word Count: 5883

Word Count (excluding abstract & references): 4440

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n.323262 to A. Mike Burton.

Abstract

First impressions formed after seeing someone's face or hearing their voice can affect many social decisions, including voting in political elections. Despite the many studies investigating the independent contribution of face and voice cues to electoral success, their integration is still not well understood. Here, we examine a novel electoral context, student representative ballots, allowing us to test the generalisability of previous studies. We also examine the independent contributions of visual, auditory and audiovisual information to social judgements of the candidates, and their relationship to election outcomes. Results showed that perceived trustworthiness was the only trait significantly related to election success. These findings contrast with previous reports on the importance of perceived competence using audio or visual cues only in the context of national political elections. The present study highlights the role of real-world context and emphasises the importance of using ecologically valid stimulus presentation in understanding real-life social judgement.

Keywords: First impressions; audiovisual integration; faces; voices; elections

Introduction

We form first impressions of unfamiliar people the moment we meet them. In such a situation we are usually presented with audio (i.e. voice) and visual (i.e. face) cues simultaneously, yet the vast majority of social evaluation literature has been focused on their independent effects (McAleer, Todorov, & Belin, 2014; Oosterhof & Todorov, 2008; Todorov, Said, Engell, & Oosterhof, 2008). First impressions from faces and voices have many parallels – they are both formed very quickly (after a 100-ms exposure for faces, Willis & Todorov, 2013 and after brief utterances for voices, McAleer et al., 2014) and have the same underlying structure, with dominance and trustworthiness emerging as fundamental dimensions (McAleer et al., 2014; Oosterhof & Todorov, 2008). While evidence for the accuracy of social judgements is limited at best (Klofstad & Anderson, 2018; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015), people seem to agree with each other's evaluations, implying they are based on some consistent visual information in the face or acoustic information in the voice (McAleer et al., 2014; Zebrowitz & Montepare, 2008). Most importantly, first impressions have been shown to influence our behaviour and decisions both in situations where appearance might be relevant, e.g. dating (Doll et al., 2014; Wells, Dunn, Sergeant, & Davies, 2009) and where we should be making more objective and informed choices such as in political elections (Ballew & Todorov, 2007; Klofstad, 2016; Klofstad, Anderson, & Peters, 2012; Olivola & Todorov, 2010; Sussman, Petkova, & Todorov, 2013; Todorov, Mandisodza, Goren, & Hall, 2005), business and finance decisions (Dean, 2017; Fruhen, Watkins, & Jones, 2015; Rule & Ambady, 2008) and court sentencing (Wilson & Rule, 2015; Zebrowitz & McDonald, 1991 also see Olivola, Funk, & Todorov, 2014; Todorov et al., 2015 for reviews).

The present paper focuses on how audio and visual cues are integrated to inform social judgements relevant to the most studied choice domain – leadership elections. Empirical work exploring the dimensions voters use when evaluating political candidates shows that competence is deemed one of the most important traits to possess (Miller, Wattenberg, & Malanchuk, 1986; Trent, Mongeau, Trent, Kendall, & Cushing, 1993). Therefore, competence has been the focus of social evaluation research on political decisions, with studies consistently demonstrating that political candidates perceived to have a more competent-looking face than their opponents are more likely to win US Senate, House of Representatives, gubernatorial and even Presidential elections (see Hall, Goren, Chaiken, & Todorov, 2009; Olivola & Todorov, 2010 for reviews). This effect has been replicated across different exposure durations (100ms, 250ms or unlimited time in Ballew and Todorov, 2007) and different decision tasks (2AFC in Ballew & Todorov, 2007 and Todorov et al., 2005, and rating the competence of multiple candidates in Sussman et al., 2013). Such findings support the assumption that first impressions represent rapid and unreflective (also referred to as ‘system 1’) judgements, which means their effect might be unnoticed by voters (Chaiken & Trope, 1999; Kahneman, 2003). In fact, one way to disturb the relationship between competence evaluations and election success is to instruct participants to make a deliberate, rather than an intuitive, decision (Ballew & Todorov, 2007).

In the voice perception literature, research on political and leadership decisions has exclusively focused on the role of vocal pitch. Tigue et al. (2012), for example, presented participants with pairs of voice recordings (one with a high and one with a low pitch) and asked them to select the person who sounded like a better leader and the one they would vote for. The results showed a significant preference for low-pitched voices both in terms of leadership and hypothetical votes. These findings have also been replicated with audio

recordings and data from the US House of Representatives elections, demonstrating a negative correlation between vocal pitch and vote share for both male and female candidates (Klofstad, 2016).

Despite all we already know about evaluating faces or voices along social dimensions, a more realistic approach to first impressions would acknowledge and explore their integration. Historically, most audiovisual literature has focused on identity and emotion recognition (Campanella & Belin, 2007; Massaro & Egan, 1996; Robertson & Schweinberger, 2010; Schweinberger, Robertson, & Kaufmann, 2007) with relatively fewer studies on social evaluation (Mileva, Tompkinson, Watt, & Burton, 2018; Rezlescu et al., 2015; Tsankova et al., 2015). These studies show that the relative importance of face and voice cues depends on the social dimension, with visual information from the face being more diagnostic of attractiveness (Rezlescu et al., 2015) and trustworthiness judgements (Mileva et al., 2018; Tsankova et al., 2015) and auditory information from the voice being more diagnostic of dominance judgements (Mileva et al., 2018; Rezlescu et al., 2015).

The relative contribution of face and voice cues to competence judgements as well as their integration in the context of leadership decisions, however, is not well understood. In one of the few studies addressing this issue, Benjamin and Shapiro (2009) showed participants 10-second video footage of political candidates in a debate. Their task was to rate each person on attractiveness, likeability, leadership and political orientation (liberal vs conservative) as well as guess which candidate won the election. The videos were presented in full sound, with muddled sound or with no sound. Their results showed that participants were able to predict the winner of the election above chance levels in all three audio conditions with no significant differences between them. However, as this was not the focus of the paper,

Benjamin and Shapiro did not provide much information about the influence of social ratings. A recent study by Klofstad (2017) also explores the relative contribution of face and voice cues to election success. However, it focuses on a single social trait – competence and on a single acoustic characteristic – vocal pitch. Here, images of House of Representatives members rated as the most and the least competent were paired with a separate set of voice recordings manipulated to have either a higher or a lower pitch. Participants were then presented with two such pairs and asked to cast a hypothetical vote. The results showed that candidates with competent faces and competent voices (i.e. voices with low pitch, see Klofstad et al., 2012) won the largest proportion of votes; however, the effect of facial competence was 2.8 times larger than the effect of vocal competence. Findings from both studies imply that visual information in the face might be of higher importance than vocal characteristics when it comes to political and leadership decisions.

An interesting question following from studies that integrate audio and visual information together is whether such cues lead to the same social evaluation. In other words, are people with trustworthy or dominant faces, also perceived to have trustworthy and dominant voices? Previous literature has been mainly focused on judgements based on a single modality rather than their integration. Attractiveness is the only exception within this context, with some evidence that people perceived as more attractive from their faces are also perceived as more attractive based on their voices (Collins & Missing, 2003; Saxton, Burriss, Murray, Rowland, & Craig, 2009; Saxton, Caryl, & Craig, 2006). There are also studies exploring the perception of physical characteristics from both face and voice cues (Puts, Jones, & DeBruine, 2012; Re, DeBruine, Jones, & Perrett, 2013). Some studies show that both visual and acoustic characteristics are highly correlated with measures of a person's strength, height and weight (Burton & Rule, 2013; Hodges-Simeon, Gurven, Puts, & Gaulin, 2014), while others report

high correlations between face- and voice-based ratings for masculinity, health and height (Smith, Dunn, Baguley, & Stacey, 2016). Together with findings of the strong relationship between facial and vocal perceived threat (Han et al., 2017), these studies suggest a possible link between dominance judgements inferred from the face and the voice.

The present study aims to extend previous literature in two ways. First, we examine a very different electoral context to that usually studied. Student representative elections are common across colleges and universities worldwide. As such, they are part of life for very large number of people. (It would be hard to estimate how many, but any estimate would presumably be in the millions, every year.) Of course, these are of no geopolitical consequence by comparison to the elections typically studied in psychology, which universally focus on political voting. As described above, the large body of research shows the importance of perceived competence in national political contests, but we do not know whether this factor will be so important in elections of all types. Candidates in student elections may attract a different type of support, perhaps based on social factors or influenced by the fact that the winners of such elections receive relatively little real power. For this reason, it is important to establish whether the influence of perceived competence is universal, or tied to a particular context.

Second, in this study we examine the independent contribution of faces, voices and their combination across different social judgements. We do this using real election material (student campaign material) and relate social judgements to real outcome (election results). The use of genuine election material, rather than hypothetical elections, allows us to examine whether differences in social judgement are powerful enough to survive the highly variable, “messy”, context of a real ballot.

Throughout the study, participants were presented with short video clips of unfamiliar candidates running for student representative elections and then asked to rate each person for the fundamental social dimensions of competence, trustworthiness, dominance and attractiveness. The experimental stimuli comprised audio cues only (voice recording extracted from each video), visual cues only (muted video clips) or audio and visual cues together (unedited clips). If the results from the political social evaluation election literature generalise across social contexts, then we would expect that competence, as judged from both the face and the voice, would be the trait most closely related to election success. Following from Klofstad (2017) we would also anticipate that visual information from the face would be more diagnostic of election success than acoustic information from the voice. However, we are interested to observe whether these patterns hold in the present context. More generally, we also predict that there will be positive correlations between face- and voice-based ratings, at least for judgements of attractiveness and dominance, where such patterns have previously been observed in neutral, lab-based settings.

Method

Participants

A total of 99 participants (7 male, mean age = 19, age range = 18-50) took part in the experiment. All were 1st year students from the University of York, who were unfamiliar with candidates from student elections held in earlier years, whose campaign material was used in the study. Participants had normal or corrected-to-normal vision and received course credit or payment for their participation. Sample size was determined with an a priori power analysis in GPower (Erdfelder, Faul, & Buchner, 1996). Sussman et al. (2013) report one of

the few studies which use a wider range of candidates and correlate the percent of votes received with their trait ratings. They collected data for 18 candidates with at least 32 participants rating each image and report a correlation of 0.53 between competence and vote percent. Based on their results, our power analysis revealed that to detect an effect of a similar size, with 90% power using an alpha of .05 (two-tailed), we need a sample of 33 participants per condition. The study was approved by the ethics committee of the Psychology Department at the University of York and informed consent was provided prior to participation.

Materials

The study used 22 videos produced by Student Television (<http://ystv.co.uk/>) as manifestos from candidates running in the University of York Student Union elections¹. We used videos of candidates running for the positions of student union president (11/22) and sports president between 2015 and 2017 (original videos can be found at <https://ystv.co.uk/watch/Elections/>)². There were 7/22 videos of female candidates. An average of 2524³ votes were cast per year for each position and winning candidates received an average of 1247 votes. Given that there were at least four candidates in each election, successful candidates won by a comfortable margin, securing about 50% of votes cast.

¹ All 60-second manifesto videos are property of YSTV, released under a Creative Commons BY-NC-SA 4.0 license.

² This is the official YSTV page with candidate manifestos from all past years and categories. Scrolling down will allow you to select a specific election year, after which you can access the 60-second manifestos for all candidates.

³ Some candidates did not record a manifesto video or recorded a manifesto that was inappropriate for the purposes of the present study. Student voters are also allowed to cast their vote for R.O.N. (Re Open Nominations) if they did not want to vote for any of the candidates. The average number of votes cast for student president and sports president between 2015-2017 increases from 2524 to 3373 when we consider all votes cast for those two positions in that time period.

All videos were cut to capture only candidates presenting themselves and the position they were running for (mean video length = 3.41s, video length range = 2s – 6s). These short clips were used in the audiovisual condition. Participants in the visual condition saw the same 22 clips presented silently, while participants in the audio condition heard the voice of the candidates extracted from the same clips.

Procedure

The study used the online platform Qualtrics (2015; Provo, UT) to collect data; however participants were tested in the lab. Each participant was randomly assigned to one of three conditions: audiovisual, video only or audio only. Participants were presented with all 22 clips and asked to rate each candidate for attractiveness, trustworthiness, dominance and competence on a 9-point Likert scale. Each trait was rated in a separate block to minimise any carryover effects (Rhodes, 2006). Block and stimulus presentation order was randomised individually for each participant.

Results

Social Traits and Election Success

All trait ratings showed good inter-rater reliability (Cronbach's α ranging from .75 to .94) and we therefore calculated an average score for each candidate within each trait x modality condition. Table 2 shows these average scores together with information about the total number and relative proportion of votes received by each candidate. The average trait scores were then correlated with the proportion of votes received by each candidate separately for the auditory, visual and audiovisual stimulus presentation (see Table 1).

Pearson correlations identified trustworthiness as the only trait related to election success. This relationship was significant in the auditory ($r = .50, p = .017, 95\% \text{ CI } [.10, .91]$), visual ($r = .45, p = .038, 95\% \text{ CI } [.03, .86]$) and in the audiovisual condition ($r = .44, p = .040, 95\% \text{ CI } [.02, .86]$). Figure 1 shows the relationship between trustworthiness ratings and proportion of votes across modality. No other trait was significantly correlated with the proportion of votes received by the candidates. In order to check the reliability of these results we also used the Benjamini-Hochberg correction (1995) for multiple comparisons with a false discovery rate of 0.2. Trustworthiness remained significantly correlated with vote proportion in all three conditions after the correction (auditory condition: $p = .017$, visual condition: $p = .033$, audiovisual condition: $p = .05$).

Table 1. Uncorrected Pearson's correlations between the proportion of votes received by each candidate and each social trait across all three presentation modalities. Significant correlations are presented in bold.

Trait	Auditory		Visual		Audiovisual	
	Presentation		Presentation		Presentation	
	Pearson's r	p	Pearson's r	p	Pearson's r	p
Attractiveness	.30	.178	.28	.214	.29	.188
Trustworthiness	.50	.017	.45	.038	.44	.040
Dominance	-.11	.613	-.16	.470	-.07	.746
Competence	.22	.334	.28	.213	.25	.266

Note. All $N_s = 22$

Table 2. Mean proportion of votes and social ratings (A – attractiveness, T – trustworthiness, D – dominance, C – competence) across the three conditions (auditory, visual and audiovisual) for each election candidate.

ID	Total Votes	Vote Proportion	A	Auditory Presentation			Visual Presentation				Audiovisual Presentation			
				T	D	C	A	T	D	C	A	T	D	C
1	1365	0.359	5.12	5.39	5.24	5.42	3.15	4.97	5.00	5.06	2.94	5.18	3.94	5.12
2	316	0.083	6.64	5.91	5.79	6.94	3.97	4.42	5.33	5.00	5.12	5.30	5.48	6.24
3	389	0.102	3.15	4.61	3.88	4.55	4.64	4.94	4.94	5.12	3.55	4.76	3.52	4.91
4	888	0.255	3.79	5.73	3.55	5.33	3.55	5.24	3.58	5.42	3.18	5.33	3.58	4.73
5	593	0.171	4.00	5.73	4.24	4.70	4.91	5.27	3.21	4.73	5.42	5.82	3.70	5.27
6	1508	0.434	6.03	5.58	4.24	5.52	6.97	6.09	5.06	7.09	7.36	6.39	5.48	7.06
7	424	0.122	5.58	5.85	6.24	6.52	5.39	5.09	5.79	6.36	5.52	5.64	5.70	6.42
8	518	0.152	5.64	5.03	5.48	5.88	4.45	5.85	3.82	5.91	3.88	5.88	3.61	5.52
9	828	0.242	4.79	5.39	5.48	5.94	4.48	5.70	4.06	6.24	4.64	5.94	4.73	6.00
10	213	0.062	5.15	4.18	6.39	5.36	4.58	4.30	5.85	5.97	4.94	4.55	6.33	6.09
11	953	0.279	5.21	5.94	3.88	5.48	4.79	5.27	4.42	5.55	4.97	6.48	4.48	6.00
12	440	0.152	5.30	4.79	4.88	4.82	6.79	5.09	5.00	5.61	6.67	5.18	4.82	5.36
13	668	0.230	5.06	4.58	4.88	5.91	4.64	4.42	4.21	4.73	5.18	4.45	4.55	5.58
14	1096	0.378	4.61	5.15	5.15	5.79	4.97	4.64	4.58	5.85	5.15	5.06	5.12	5.70
15	599	0.207	6.03	5.61	6.39	6.67	4.61	5.00	4.85	5.39	5.73	6.06	5.30	6.64
16	275	0.082	5.00	5.67	5.24	6.03	4.94	6.55	4.24	6.73	4.97	7.00	4.82	6.45
17	1146	0.343	5.61	5.58	4.64	5.79	3.33	4.73	3.94	4.06	3.85	5.24	3.15	4.73
18	113	0.034	3.42	4.85	5.09	5.33	3.24	3.82	4.39	4.76	2.94	4.03	4.36	5.00
19	50	0.015	3.88	3.91	4.61	4.70	2.88	4.12	4.97	4.82	3.12	4.33	5.33	5.15
20	807	0.244	5.61	6.30	4.61	5.79	5.18	6.24	5.12	6.73	5.97	7.00	4.24	6.67
21	1413	0.428	5.21	6.27	6.36	6.52	5.33	6.73	4.88	6.85	5.55	7.15	5.88	6.94
22	542	0.164	4.70	5.09	4.82	4.91	4.91	5.09	5.18	5.15	5.21	5.61	4.18	5.36

Note. Vote proportion reflects the proportion of votes that each candidate received in their relative race. Some candidates did not record a manifesto video, however, the votes they received have been taken into consideration when calculating the proportion of votes received by the present candidates.

Although participants rated candidates on each trait in separate blocks with randomised order, there were only 22 candidates which could still lead to some potential carryover effects (Rhodes, 2006). In order to address this issue, we used data from the first rating block for each participant and used ratings across all conditions as they present with a very similar pattern of results. This resulted in 23 participants rating attractiveness, 30 participants rating trustworthiness, 22 participants rating dominance and 24 participants rating competence. Consistent with our earlier findings, Pearson correlations showed a significant relationship between vote proportion and trustworthiness ($r = .67, p = .001, 95\% \text{ CI } [.33, 1.02]$). No other traits were significantly correlated with the proportion of votes received by the candidates.

Relationship Between Face and Voice Cues

Such findings imply that the effect of face and voice cues might be complementary, rather than independent. To explore this further, we looked at the correlations between ratings attributed to each candidate when participants were presented with auditory or visual cues only (see Figure 2). Pearson correlations showed a positive relationship between ratings attributed to faces and voices for all social traits. These correlations were strongest for ratings of trustworthiness ($r = .63, p = .002, 95\% \text{ CI } [.27, .99]$) and dominance ($r = .47, p = .028, 95\% \text{ CI } [.06, .88]$), demonstrating that people who are perceived as more trustworthy and dominant as judged from their faces, receive similar ratings based on their voices. The correlations between face and voice ratings for attractiveness ($r = .36, p = .105, 95\% \text{ CI } [-.08, .79]$) and competence ($r = .33, p = .132, 95\% \text{ CI } [-.11, .77]$) also followed the same direction but were not significant.

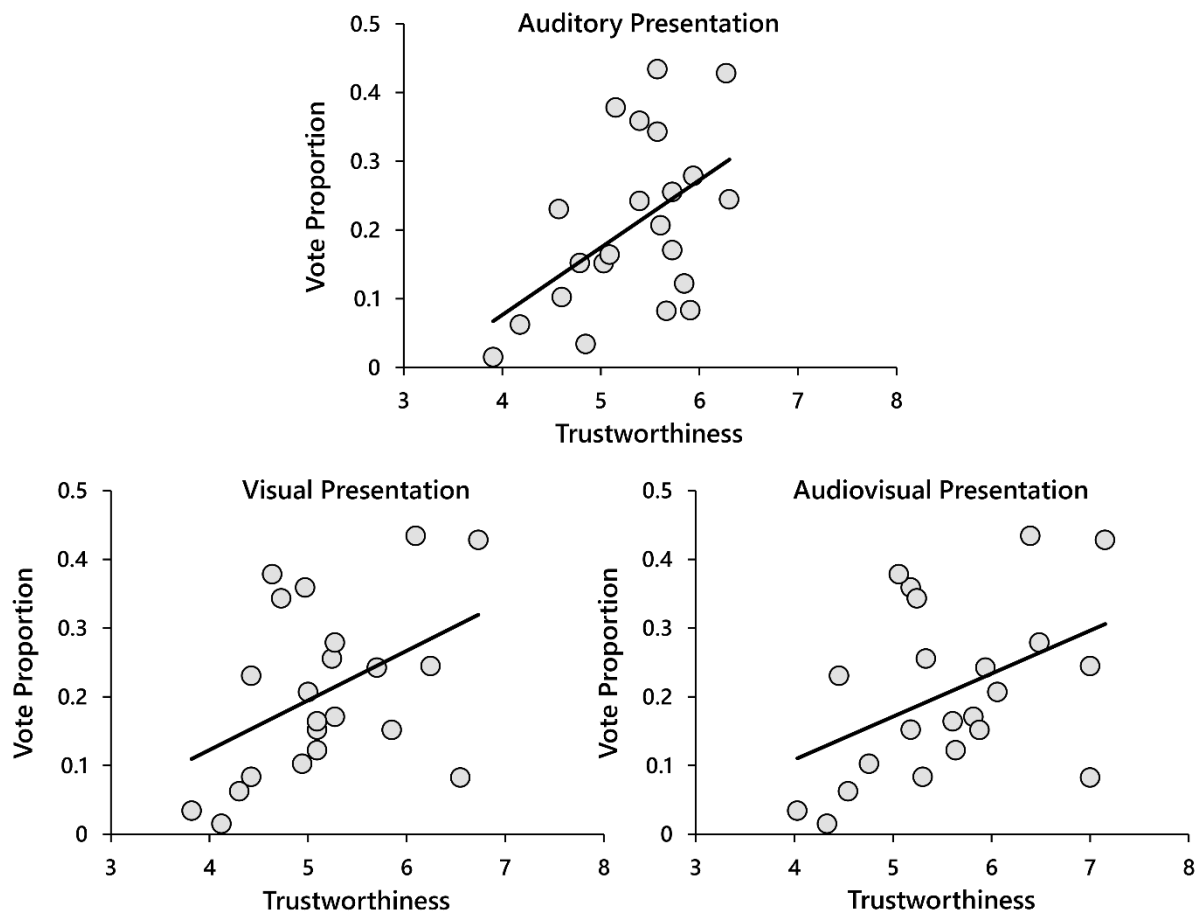


Figure 1. Correlations between trustworthiness and proportion of votes in the auditory only (top), visual only (bottom left) and audiovisual (bottom right) conditions.

Acoustic Measures

Since trustworthiness as judged from the voice was significantly correlated with election success, we extracted a number of acoustic measures using the ProsodyPro script (Xu, 2013) in Praat (Boersma and Weenink, 2016). The acoustic parameters included: 1) speech rate, calculated as the average number of syllables produced per second of speech, as a measure of how quickly each utterance was produced; 2) fundamental frequency (F0) range, as a measure of how much variation in intonation was present; 3) median F0 as an average measure of how high-pitched a speaker's voice was (a measure preferable to mean F0, as it

reduces influence from outliers due to octave jumps, Lindh, 2006); 4) mean intensity; 5) formant dispersion between F1 and F3, calculated as the average distance between the first three formants; 6) vocal jitter, measured as the “mean absolute difference between consecutive periods, divided by [the]mean period” (Xu, 2013); 7) vocal shimmer, measured as the “mean absolute difference between amplitudes of consecutive periods, divided by the mean amplitude” (Xu, 2013); 8) harmonic-to-noise ratio, as a measure of the “degree of acoustic periodicity” (Xu, 2013). For male speakers, the pitch calculation range in Praat was set between 75Hz-300Hz, whereas for female speakers the range was set at 100-500Hz. These values conform to the normative values recommended by Boersma and Weenink (2016). No single acoustic cue in the voice acted as a consistently reliable predictor of voting behaviour, implying that participants were using other cues to inform their social judgements.

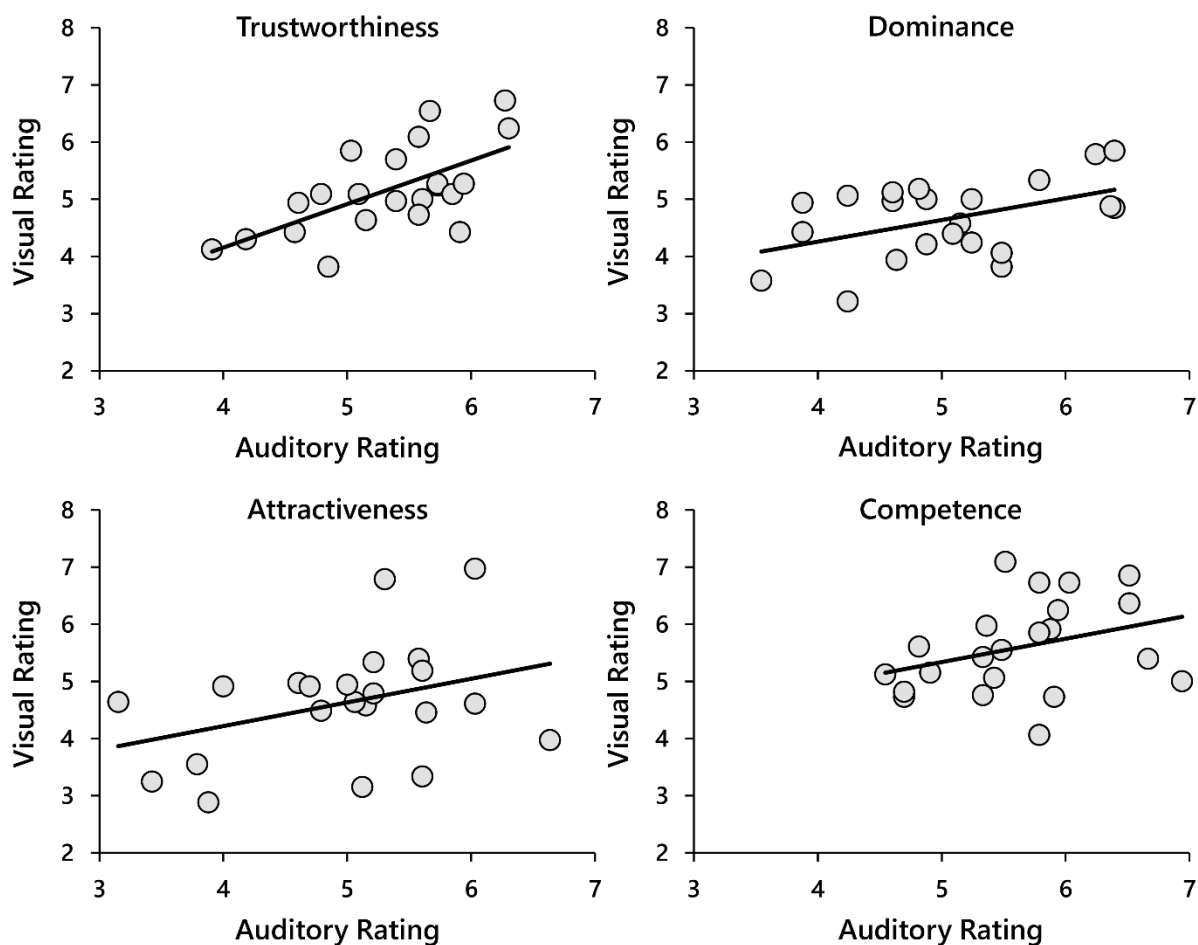


Figure 2. Relationships between ratings based on auditory and visual cues in the perception of trustworthiness (top left), dominance (top right), attractiveness (bottom left) and competence (bottom right).

Discussion

The present paper aimed to explore the relative contribution of auditory and visual cues to social traits associated with success in a novel context: student elections. Participants rated student representative campaign videos capturing candidates' own faces and voices on a number of social dimensions. These ratings were then correlated with the proportion of votes received by each candidate. Our findings showed that trustworthiness was the only trait related to the election outcome. While this was true in all three modalities (auditory, visual and audiovisual), trustworthiness as judged from the candidates' faces was the best predictor of election success.

These results are particularly interesting given previous data on the role of perceived competence in election and leadership decisions. It is possible that these different findings reflect the use of real, dynamic (rather than photographic) stimuli. However, it seems highly likely that they are influenced by the different social contexts, for example, electing a President of the USA is rather unlike electing a student sports representative. Indeed, there is already evidence that the context of an election can guide the dimensions people use when making their decisions. Little, Burriss, Jones and Roberts (2007), for example, collected social ratings based on the unrecognisable morphed images of George W. Bush and John Kerry and asked participants to cast hypothetical votes in two different contexts – a time of

war and a time of peace. The results showed a strong preference for the morphed face of Bush in time of war but the morphed face of Kerry received a higher proportion of votes in time of peace. Critically, Bush's morphed face was perceived as more dominant and masculine, whereas Kerry was seen as more likeable and intelligent. There is also evidence for changes in the importance of assigned to each social trait in different cultures and different countries. Berggren, Jordahl and Poutvaara (2010), for example, show that ratings of attractiveness are a better predictor of election success in Finland, while Rule et al. (2010) report that judgements of warmth predict elections outcome the best for Japanese participants.

Our findings on the importance of face cues for election decisions are in line with those of Klofstad (2017) as well as with previous findings on the greater contribution of face cues when judging trustworthiness (Mileva et al., 2018; Tsankova et al., 2015). It should, however, be noted that trustworthiness as judged from the face and from the voice were very highly correlated, suggesting that first impressions from faces and voices both signal the same integrated person evaluation. This is further supported by the highly positive correlations between face- and voice-based ratings of trustworthiness and dominance as well those for ratings of attractiveness and competence, although they were not significant. It is therefore possible that the effects of face and voice cues in social evaluation are complementary, rather than independent. This is an important finding as most previous research has been unable to address this issue. Most studies have used face and voice stimuli of different identities paired together or manipulated voices artificially (Klofstad, 2017; Mileva et al., 2018) instead of the ecologically valid approach we adopt here. It should, however, be noted that our analysis was based on a relatively small stimulus sample which could potentially affect the reliability and generalisability of our findings. Nevertheless, we report a very consistent pattern of results

across all three presentation conditions which helps strengthen our interpretation and conclusions.

Overall, this study shows that trustworthiness emerges as the most important trait in student elections when information about candidates' own faces and voices is available. Our results support the role of context in the selection of social traits associated with electoral success. Most importantly, given that this is the first study to integrate face and voice cues in a more ecologically valid way, its findings provide a more complete account of the role of first impressions in predicting the outcome of leadership decisions.

References

- Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948–17953. <https://doi.org/10.1073/pnas.0705435104>
- Benjamin, D. J., & Shapiro, J. M. (2009). Thin-slice forecasts of gubernatorial elections. *Review of Economics and Statistics*, *91*(3), 523–536. <https://doi.org/10.1162/rest.91.3.523>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300. <https://doi.org/10.2307/2346101>
- Berggren, N., Jordahl, H., & Poutvaara, P. (2010). The looks of a winner: Beauty and electoral success. *Journal of Public Economics*, *94*(1–2), 8–15. <https://doi.org/10.1016/j.jpubeco.2009.11.002>
- Boersma, P., & Weenink, D. (2016). Praat: Doing phonetics by computer (Version 6.0.22) [Computer Program]. Retrieved from <http://www.fon.hum.uva.nl/praat/>
- Burton, C. M., & Rule, N. O. (2013). Judgments of height from faces are informed by dominance and facial maturity. *Social Cognition*, *31*(6), 672–685. <https://doi.org/10.1521/soco.2013.31.6.672>
- Campanella, S., & Belin, P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, *11*(12), 535–543. <https://doi.org/10.1016/j.tics.2007.10.001>
- Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: Guilford Press.

- Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour*, *65*(5), 997–1004. <https://doi.org/10.1006/anbe.2003.2123>
- Dean, D. H. (2017). The benefit of a trustworthy face to a financial services provider. *Journal of Services Marketing*, *31*(7), 771–783. <https://doi.org/10.1108/JSM-06-2016-0221>
- Doll, L. M., Hill, A. K., Rotella, M. A., Cárdenas, R. A., Welling, L. L. M., Wheatley, J. R., & Puts, D. A. (2014). How well do men's faces and voices index mate quality and dominance? *Human Nature*, *25*(2), 200–212. <https://doi.org/10.1007/s12110-014-9194-3>
- Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 1–11. <https://doi.org/10.3758/BF03203630>
- Fruhen, L. S., Watkins, C. D., & Jones, B. C. (2015). Perceptions of facial dominance, trustworthiness and attractiveness predict managerial pay awards in experimental tasks. *Leadership Quarterly*, *26*(6), 1005–1016. <https://doi.org/10.1016/j.leaqua.2015.07.001>
- Hall, C. C., Goren, A., Chaiken, S., & Todorov, A. (2009). Shallow cues with deep effects: Trait judgements from faces and voting decisions. In E. Borgida, C. M. Federico, & J. L. Sullivan (Eds.), *The political psychology of democratic citizenship* (pp. 583–605). New York: Oxford University Press.
- Han, C., Kandrik, M., Hahn, A. C., Fisher, C. I., Feinberg, D. R., Holzleitner, I. J., ... Jones, B. C. (2017). Interrelationships among men's threat potential, facial dominance, and vocal dominance. *Evolutionary Psychology*, *15*(1), 1–4. <https://doi.org/10.1177/1474704917697332>
- Hodges-Simeon, C. R., Gurven, M., Puts, D. A., & Gaulin, S. J. C. (2014). Vocal fundamental and formant frequencies are honest signals of threat potential in

peripubertal males. *Behavioral Ecology*, 25(4), 984–988.

<https://doi.org/10.1093/beheco/aru081>

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality.

American Psychologist, 58(9), 697–720. <https://doi.org/10.1037/0003-066X.58.9.697>

Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698–2704.

<https://doi.org/10.1098/rspb.2012.0311>

Klofstad, Casey A. (2016). Candidate voice pitch influences election outcomes. *Political*

Psychology, 37(5), 725–738. <https://doi.org/10.1111/pops.12280>

Klofstad, Casey A. (2017). Looks and sounds like a winner: Perceptions of competence in candidates' faces and voices influences vote choice. *Journal of Experimental Political Science*, 4(03), 229–240. <https://doi.org/10.1017/XPS.2017.19>

Klofstad, Casey A., & Anderson, R. C. (2018). Voice pitch predicts electability, but does not signal leadership ability. *Evolution and Human Behavior*, 39(3), 349–354.

<https://doi.org/10.1016/j.evolhumbehav.2018.02.007>

Lindh, J. (2006). Preliminary F0 statistics and forensic phonetics. *Proceedings of the 15th Annual International Association of Forensic Phonetics and Acoustics Conference. Sweden: Department of Linguistics, Göteborg University.*

Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28(1), 18–27.

<https://doi.org/10.1016/j.evolhumbehav.2006.09.002>

Massaro, D. W., & Egan, P. B. (1996). Perceiving affect from the voice and the face.

Psychonomic Bulletin and Review, 3(2), 215–221. <https://doi.org/10.3758/BF03212421>

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say “hello”? Personality

impressions from brief novel voices. *PLoS ONE*, 9(3), 1–9.

<https://doi.org/10.1371/journal.pone.0090779>

Mileva, M., Tompkinson, J. A., Watt, D., & Burton, A. M. (2018). Audiovisual integration in

social evaluation. *Journal of Experimental Psychology: Human Perception and*

Performance., 44(1), 128–138. <https://doi.org/https://doi.org/10.1037/xhp0000439>

Miller, A. H., Wattenberg, M. P., & Malanchuk, O. (1986). Schematic assessments of

presidential candidates. *American Political Science Review*, 80(2), 521–540.

<https://doi.org/10.2307/1958272>

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human

choices. *Trends in Cognitive Sciences*, 18(11), 566–570.

<https://doi.org/10.1016/j.tics.2014.09.007>

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait

inferences and voting. *Journal of Nonverbal Behavior*, 34(2), 83–110.

<https://doi.org/10.1007/s10919-009-0082-1>

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings*

of the National Academy of Sciences, 105(32), 11087–11092.

<https://doi.org/10.1073/pnas.0805664105>

Puts, D. A., Jones, B. C., & DeBruine, L. M. (2012). Sexual selection on human faces and

voices. *Journal of Sex Research*, 49(2–3), 227–243.

<https://doi.org/10.1080/00224499.2012.658924>

Re, D. E., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2013). Facial cues to perceived

- height influence leadership choices in simulated war and peace contexts. *Evolutionary Psychology*, *11*(1), 89–103. <https://doi.org/10.1177/147470491301100109>
- Rezlescu, C., Penton, T., Walsh, V., Tsujimura, H., Scott, S. K., & Banissy, M. J. (2015). Dominant voices and attractive faces: The contribution of visual and auditory information to integrated person impressions. *Journal of Nonverbal Behavior*, *39*(4), 355–370. <https://doi.org/10.1007/s10919-015-0214-8>
- Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*(1), 199–226. <https://doi.org/10.1146/annurev.psych.57.102904.190208>
- Robertson, D. M. C., & Schweinberger, S. R. (2010). The role of audiovisual asynchrony in person recognition. *Quarterly Journal of Experimental Psychology*, *63*(1), 23–30. <https://doi.org/10.1080/17470210903144376>
- Rule, N. O., & Ambady, N. (2008). The face of success. *Psychological Science*, *48*, 109–111. <https://doi.org/10.1111/j.1467-9280.2008.02054.x>
- Rule, N. O., Ambady, N., Adams, R. B., Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology*, *98*(1), 1–15. <https://doi.org/10.1037/a0017673>
- Saxton, T. K., Burriss, R. P., Murray, A. K., Rowland, H. M., & Craig, R. S. (2009). Face, body and speech cues independently predict judgments of attractiveness. *Journal of Evolutionary Psychology*, *7*(1), 23–35. <https://doi.org/10.1556/JEP.7.2009.1.4>
- Saxton, T. K., Caryl, P. G., & Craig, R. S. (2006). Vocal and facial attractiveness judgments of children, adolescents and adults: The ontogeny of mate choice. *Ethology*, *112*(12), 1179–1185. <https://doi.org/10.1111/j.1439-0310.2006.01278.x>
- Schweinberger, S. R., Robertson, D., & Kaufmann, J. M. (2007). Hearing facial identities.

Quarterly Journal of Experimental Psychology, 60(10), 1446–1456.

<https://doi.org/10.1080/17470210601063589>

Smith, H. M. J., Dunn, A. K., Baguley, T., & Stacey, P. C. (2016). Concordant cues in faces and voices: Testing the Backup Signal Hypothesis. *Evolutionary Psychology*, 14(1), 1–10. <https://doi.org/10.1177/1474704916630317>

Sussman, A. B., Petkova, K., & Todorov, A. (2013). Competence ratings in US predict presidential election outcomes in Bulgaria. *Journal of Experimental Social Psychology*, 49(4), 771–775. <https://doi.org/10.1016/j.jesp.2013.02.003>

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210–216. <https://doi.org/10.1016/j.evolhumbehav.2011.09.004>

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308(5728), 1623–1626. <https://doi.org/10.1126/science.1110589>

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, 66(1), 519–545. <https://doi.org/10.1146/annurev-psych-113011-143831>

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, 12(12), 455–460. <https://doi.org/10.1037/a0027950>

Trent, J. S., Mongeau, P. A., Trent, J. D., Kendall, K. E., & Cushing, R. B. (1993). The ideal candidate: A study of the desired attributes of the public and the media across two

presidential campaigns. *American Behavioral Scientist*, 37(2), 225–239.

<https://doi.org/https://doi.org/10.1177/0002764293037002010>

Tsankova, E., Krumhuber, E., Aubrey, A. J., Kappas, A., Möllering, G., & Rosin, P. L.

(2015). The multi-modal nature of trustworthiness perception (pp. 147–152).

Proceedings of the First Joint Conference on Facial Analysis and Auditory-Visual

Speech Processing. Retrieved from [http://www.isca-](http://www.isca-speech.org/archive/avsp15/index.html)

[speech.org/archive/avsp15/index.html](http://www.isca-speech.org/archive/avsp15/index.html)

Wells, T. J., Dunn, A. K., Sergeant, M. J. T., & Davies, M. N. O. (2009). Multiple signals in

human mate selection: A review and framework for integrating facial and vocal signals.

Journal of Evolutionary Psychology, 7(2), 111–139.

<https://doi.org/10.1556/JEP.7.2009.2.2>

Willis, J., & Todorov, A. T. (2013). Making up your mind after 100-ms exposure to face.

Psychological Science, 17(7), 592–598. [https://doi.org/10.1111/j.1467-](https://doi.org/10.1111/j.1467-9280.2006.01750.x)

[9280.2006.01750.x](https://doi.org/10.1111/j.1467-9280.2006.01750.x)

Wilson, J. P., & Rule, N. O. (2015). Hypothetical sentencing decisions are associated with

actual capital punishment outcomes: The role of facial trustworthiness. *Social*

Psychological and Personality Science, 7(4), 331–338.

<https://doi.org/10.1177/1948550615624142>

Xu, Y. (2013). ProsodyPro - A tool for large-scale systematic prosody analysis. *Proceedings*

of Tools and Resources for the Analysis of Speech Prosody, 7–10.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and

attractiveness on adjudications in small claims courts. *Law and Human Behavior*, 15(6),

603–623. <https://doi.org/10.1007/BF01065855>

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, 23(10), 1497–1517.

<https://doi.org/10.1111/j.1751-9004.2008.00109.x>