More than words: Judgments of politicians and the role of different communication channels

Markus Koppensteiner *, Pia Stephan, Johannes Paul Michael Jäschke

University of Vienna, Department of Anthropology, Austria

**A R T I C L E  I N F O**

**Article history:**
Received 2 April 2015
Revised 9 May 2015
Accepted 12 May 2015
Available online 30 June 2015

**Keywords:**
Nonverbal communication
Body movements
Snap judgments
Personality
Thin slices
Impression formation

**A B S T R A C T**

We collected short video clips of speakers and created five types of stimuli: (1) the original videos, (2) the audio tracks only, (3) single pictures only, (4) speech content, and (5) stick-figure animations displaying body motion. Participants rated these stimuli on a brief Big Five personality inventory. We then used ratings of the incomplete information conditions to predict ratings of the original video condition. Impressions in the audio track condition were strong predictors throughout all trait ratings. However, other cues were also non-negligible contributors to an overall impression. People even make sense of parsimonious cues, e.g., an animated stick-figure. Thus, presenters on a public stage are not only judged by what they say but also by how they move.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

**1. Introduction**

When people form impressions about others, words often seem to affect them less than the observed outward appearance and nonverbal behavior. Visual and auditory cues affect people's judgments of their interaction partners; judgments that are made spontaneously, effortlessly, and without conscious processing (Ambady, Bernieri, & Richeson, 2000; Sunnafrank & Ramirez, 2004). On the one hand, first impressions can be misleading and a source of prejudices and stereotyping (Zebrowitz & Montepare, 2005). On the other hand, they can pick up relevant information about one's social environment. After being exposed to brief extracts of nonverbal or verbal information, people are able to assess other people's actual personality traits, their job performances, or a CEO's abilities to generate company profits (Albright, Kenny, & Malloy, 1988; Ambady et al., 2000; Borkenau & Liebler, 1992a,b; Hecht & LaFrance, 1995; Rule & Ambady, 2008; Scherer, 1978).

Irrespective of whether they are the key to someone's actual personality traits or abilities, snap judgments can have a strong impact on impression formation and decision making. This is greatly important for those who enter the public arena. Politicians and leaders who vie for media attention and try to win the approval of an audience have to be aware that they are not only judged by the content they present. Nonverbal and salient cues are assumed to be processed efficiently and easily remembered and for this reason they can dominate over verbal information (Clark & Pavio, 1991) Their impact may even be more prevailing nowadays because news reports have been undergoing a shift from political content to image bites (see Bucy & Grabe, 2007; Stewart, 2010). Moreover, the flood of information that people are confronted with daily imposes an additional cognitive load, which increases the propensity to take mental shortcuts when making decisions (Olivola & Todorov, 2010). Thus, people may tend to choose their leaders not after careful deliberation but on the basis of superficialities.

Empirical studies underscore that appearance cues and nonverbal behaviors influence judgments of politicians and other leaders. Research on the perception of charisma showed that potential leaders who display expressive non-verbal behaviors (i.e., more body gestures, more variations in intonation, more eye-contact, etc.) are seen as more charismatic than persons who showed less non-verbal behaviors (Awamleh & Gardner, 1999; Gardner, 2003; Holladay & Coombs, 1993, 1994). Moreover, experiments using manipulated voices revealed that vocal cues affect people's attributions of leadership qualities and their voting behavior (Klofstad, Anderson, & Peters, 2012; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012). People even read leadership qualities, such as, competence, trustworthiness, or dominance, into photographs of political candidates. Interestingly, the consensus among such ratings is strong enough to make them reliable predictors of hypothetical voting decisions and actual election outcomes.
All these empirical findings point to nonverbal cues as the prevailing influence on the perception of leaders and politicians. However, some researchers who compared the relative impact of visual, vocal, and verbal information on judgments of politicians came to different conclusions (Krauss, Apple, Morency, Wenzel, & Winton, 1981; Nagel, Maurer, & Reinemann, 2012). They found that speech content dominates over nonverbal information. This does not undermine the role of nonverbal behaviors in human communication but indicates that their influence varies with the situation in which behaviors are performed as well as with audience motivation and involvement (Allwood, 2002; O'Sullivan, Ekman, Friesen, & Scherer, 1985; Petty, Cacioppo, & Schumann, 1983).

In summary, regardless of on which communication channel (i.e., vocal, visual, or verbal) people's first impressions are based when making inferences of social relevance, research clearly shows that “thin slices” of behavior or appearance cues can have a strong impact on which traits and abilities people read into their social environment.

1.1. Motion cues as social information

Body motion is a form of nonverbal channel that comprises hand gestures, movements of the head, or position shifts of the whole body. Although the current study provided data on the interplay of different communication channels in impression formation we mainly focus on the relative role of body motion.

Empirical studies have found that people are very adept at extracting information from motion cues. Even abstract stimuli such as circles or triangles flitting around on a screen, are often interpreted as animal or human behavior and elicit attributions of intentionality and personality (Heider & Simmel, 1944; Koppensteiner, 2011; Scholl & Tremoulet, 2000). Research on human body motion was strongly influenced by the point light approach, which was introduced by Gunnar Johansson (1973). He and his colleagues attached point lights and reflective markers to persons’ major joints and filmed them so that only a set of dots were visible in the resulting movies. Supporting the role of motion in human perception, these kinds of stimuli only became human-like when the dots were moving, whereas observers who saw single pictures of the movies perceived nothing but a random distribution of dots. Other researchers were inspired by Johansson’s methodological approach and demonstrated that such point light displays contain enough information to make quite accurate guesses of other people's age and sex (Dittrich, Troscianko, Lea, & Morgan, 1996; Kozlowski & Cutting, 1978; Montepare & Zebrowitz-McArthur, 1988; Troje, 2002).

Research using more elaborated versions of this technique or alternative methods of motion capture revealed that motion cues convey information of social relevance (Blake & Shiffrar, 2007; Chouchourelou, Matsuka, Harber, & Shiffrar, 2006). People appear to be able to perceive affect in arm movements (Polllick, Paterson, Brunderlin, & Sanford, 2001), emotions in movements of the whole body (Atkinson, Dittrich, Gemmell, & Young, 2004; Clarke, Bradshaw, Field, Hampson, & Rose, 2005), and personality in patterns of human gait (Thoresen, Vuong, & Atkinson, 2012). The frequency and duration of motion and other kinematic features play a role in mating behavior (Bente, Donaghy, & Suwelack, 1998; Grammer, Honda, Juette, & Schmitt, 1999) and affect the way females judge the attractiveness of male dancers (Neave et al., 2011). Moreover, self-ratings and observer-ratings of personality on scales measuring sensation seeking or the Big Five personality dimensions are related to the motion behavior of dancers (Bechinie & Grammer, 2003; Hugill, Fink, Neave, Besson, & Bunse, 2011; Luck, Saarikallio, Burger, Thompson, & Toivainen, 2010). In the domain of politics, variations in body motion influence how people judge the personality and health of politicians (Kempter, 1998; Koppensteiner, 2013; Koppensteiner & Grammer, 2010; Kramer, Arend, & Ward, 2010).

Such results clearly show that motion cues are an equally important nonverbal communication channel as appearance cues (e.g., clothes), vocal cues (e.g., voice pitch) and facial expressions. The speed, the duration, and the flow of a gesture or variations in the movements of the whole body seem to have a non-negligible impact on how people form first impressions of their social environment.

1.2. The present study

Human communication works on different levels ranging from symbolic information mostly conveyed by verbal content to information that has no definite signal character and is expressed by certain qualities of motion. The studies of Mehrabian (1972) and those on the perception of charisma (see above) provide evidence that under some experimental conditions, visual information exerts a dominant influence in impression formation followed by voice quality and speech content on the last position. In addition, comparison of trait judgments based on full channel information (i.e., video with speech) with trait judgments based on incomplete information only (e.g., silent videos or voice only) indicates that cues from different communication channels convey redundant information (Borkenau & Liebler, 1992a; Friedman, Oltmanns, Gleason, & Turkheimer, 2006). However, there are variations. Some traits appear to be preferably ascribed to visual cues, while other traits are preferably ascribed to cues from other modalities (Friedman et al., 2006; Gifford, 1994; Naumann, Vazire, Rentfrow, & Gosling, 2009; Zebrowitz-McArthur & Montepare, 1989). Moreover, when asked to identify emotional expressions or statements of agreement and disagreement in political debates people are more accurate when multimodal information is available (Bänziger, Mortillaro, & Scherer, 2012; Mehu & van der Maaten, 2014).

In this study we examined to what degree body motion affects social judgments relative to information from other verbal and nonverbal communication channels. To accomplish this, we broke down short video clips of politicians making a speech into five different versions of stimuli. Independent samples of participants then judged either the full channel version of the speeches, sound only, single pictures taken out of the speeches, the content of the speeches read by a computer voice, or stick-figure animations displaying the body movements of the speakers. Measures of the participants’ first impressions were obtained using a brief questionnaire that builds on the five-factor model of personality (i.e. Big Five), because the five-factor model had already been successfully applied in numerous “thin slices” studies (e.g., Borkenau & Liebler, 1992a,b; Friedman et al., 2006). These measures allowed us to estimate to what extent information from different communication channels influences snap judgments of the speakers and how different “portions” of nonverbal and verbal information are related to body motion.

Our study extends previous research in several aspects. First, other studies on nonverbal communication often instruct actors to display specific behaviors. In contrast to that, the stimuli we used were real politicians that had given their speeches in the German parliament. Hence, ecological validity of the displayed behaviors is high. Second, although our experimental design was inspired by previous studies on the role of different communication channels (Borkenau & Liebler, 1992a,b; Friedman et al., 2006), these studies did not examine the behavior of speakers in...
a public arena. Finally, we applied other tools for stimulus preparation and stimulus presentation than other studies in the field. By translating the body movements of the speakers into animated stick figures, we diminished the influence of confounding variables and were able to determine for which traits body motion was a strong predictor and to what degree it interacts with information from other nonverbal and verbal sources. In previous work we already used stick-figure animations as stimuli and related data-driven descriptors of body motion (e.g., measures of amplitude height) to judgments of personality. This revealed that people form first impressions on the basis of simple cues embedded in the behavioral stream. However, it was unclear whether and to what extent body motion affects judgements of speakers when vocal and other cues are also available to observers (Koppensteiner & Grammer, 2010). The work presented here, which is based on the same method of motion capturing but on a new set of stimuli, is a first step to overcome such limitations.

It is suggested that rapidly evaluating other people’s intentions, assessing one’s social environment on broad trait categories and making fast decisions on whether to approach or avoid an individual has an adaptive function (e.g., Buss & Greiling, 1999; Oosterhof & Todorov, 2008). Misjudgments can have negative consequences and for this reason people may base their social evaluations on cues from different communication channels because this makes evaluations more reliable. Previous work has already shown that first impressions of extraversion, agreeableness and emotional stability are related to motion cues and gesturing (Borkenau & Liebler, 1992b; Gifford, 1994; Koppensteiner, 2013; Koppensteiner & Grammer, 2010). Hence, we hypothesized that for these personality dimensions judgments of the speakers’ body movements (i.e., stick-figure animations) can serve as predictors of people’s judgments in the full channel condition (i.e., video plus sound). Moreover, we expected pronounced links between ratings of the voice only condition and ratings of body motion, because it is well established that gesturing accompanies speech (e.g., McNeill, 1985).

Appearance cues and vocal cues have been found to be related to all dimensions of the Big-Five (Borkenau & Liebler, 1992b; Friedman et al., 2006; Naumann et al., 2009). Consequently, we expected to replicate such findings in our study. Research on nonverbal communication revealed speech content to play a minor role when people form first impressions (e.g., Awamleh & Gardner, 1999). However, there seems to be a link between conscientiousness and speech content (e.g., Friedman et al., 2006), which we also expected to reveal in our data.

2. Method

2.1. Stimulus preparation and experimental conditions

Sixty speeches taking place in the German parliament (30 male and 30 female speakers) were randomly selected out of three parliamentary sessions (November 29 & 30 and December 14, 2012). The selected speakers were non-prominent members of the parliament and were unknown to our participants. At random, we extracted video segments with a length of 15 s from these speeches. Afterwards, the video segments were used to create five different stimulus types for the experimental conditions described below.

2.1.1. Full channel condition

For the full channel condition we merely removed the lower portion of the video clip, which gave information about the speaker’s name and party affiliation. Consequently, in this experimental condition the participants of a rating experiment assessed the original video clips, which included all visual and all auditory information.

2.1.2. Static visual condition

To eliminate information on body movement we chose one picture out of the frame series of each video segment. Each of these pictures showed the speakers adopting similar body postures. For more detail, we chose pictures with the speakers standing upright having their arms positioned at their sides. In accordance with the length of all other stimulus types pictures were shown for 15 s in the rating experiment.

2.1.3. Original voice condition

For this experimental condition, we extracted the audio track from each original video segment. Thus, visual information was removed and the participants assessing such stimuli only listened to what was being said.

2.1.4. Artificial voice condition

We extracted what said during the selected video segments from the transcripts of each parliamentary session. These transcripts were converted into audio files using the document reader software Ghostreader. Because all transcripts were read aloud by the same computer voice during the rating experiment, they served as the control condition for speech content and wording.

2.1.5. Stick figure condition

The main focus of the study was on the analysis of body motion in relation to information presented in the other conditions. To this end we created stimuli, which display body motion without confounding information, e.g., appearance cues.

To accomplish this, we used a program that allows running through the video clips step by step. In the first frame of each video clip, we positioned landmarks on the speaker’s forehead, the hollow of the throat between the collar bones, ears, shoulders, elbows, hands, a spot in the middle of the body near the navel, and the corners of the lectern. Motion behavior occurring between single frames of the video clips was traced by rearranging the landmarks with the computer-mouse and software routines that automatically tracks positions shifts (Koppensteiner, 2013). Based on this data, we created stick figure animations, which served as abstract representations of the speakers’ body movements. To reduce the workload during the encoding process, we only used every third frame and filled in missing frames by linear interpolation.

2.2. Participants

A total of 308 Caucasian participants (i.e., students of the Faculty of Life Sciences) was recruited at the University of Vienna for taking part in the rating experiments. Sixty-five persons (37 females and 28 males; M-age = 23.3 years, SD = 5.8) participated in the full channel condition, 60 people (35 females and 25 males; M-age = 22.8 years, SD = 3.4) in the static visual condition, 63 (33 females and 30 males; M-age = 23.8 years, SD = 4.4) in the original voice condition, 60 participants (35 females and 25 males; M-age = 23.4 years, SD = 4.5) in the artificial voice condition, and 60 persons (33 females and 27 males; M-age = 22.5 years, SD = 3.7) in the stick figure condition. Participants received a financial compensation of €5.

2.3. Procedure

Participants were approached in person and asked to take part in a short rating experiment. After accompanying them to our laboratory, we informed the participants briefly about the
experimental set-up and the task to be accomplished. Subsequently, participants performed the rating tasks on their own using a computer-controlled interface. They were instructed to rate the stimuli after they had watched or listened to them. Pictures in the static visual condition were shown for 15 s, but participants were told that they could start their ratings whenever they felt to be ready for them. In the full channel condition, the original voice condition, and the artificial voice condition, participants wore headphones (AKG K 272 HD) connected to the computer to optimize sound quality. Stimuli were presented on the left-hand side of the user interface, while rating scales were displayed on the right-hand side. Participants completed their ratings by dragging a slider to the favored position between the right pole (i.e., named, strongly disagree) and the left pole (i.e., named, strongly agree) of the scale using a computer mouse. The slider position corresponded to a position on a scale divided into 200 subunits. All ratings started with the slider being in the neutral position (i.e., middle position on the slider bar).

To examine personality ratings of the speakers, we used the German version (Muck, Hell, & Gosling, 2007) of Gosling, Rentfrow, & Swann’s (2003) Ten Item Personality Measure (TIPI). This questionnaire is based on the five-factor model of personality and covers personality dimensions openness, extraversion, conscientiousness, emotional stability, and agreeableness. In each experimental condition, each participant rated a subset of 20 stimuli that were randomly selected from the 60 stimuli available.

### 2.4. Statistical analysis

Trait ratings were averaged for each condition and each stimulus. This yielded a dataset containing 60 speakers that were rated on ten items in five experimental conditions. Corresponding items of the TIPI questionnaire were turned into the Big Five personality dimensions according to the instructions by Gosling et al. (2003). Because each personality dimension only comprised two items, internal consistency of the questionnaire was determined by calculating the Spearman–Brown coefficient (Eisinga, Te Grotenhuis, & Pelzer, 2013).

To estimate the relative influence of voice, speech content, appearance, and body motion on judgments of personality, we correlated ratings collected in the experimental conditions presenting incomplete information with ratings of the full channel condition (i.e., original video clips). We also performed multiple regression analyses using ratings in the full channel condition as criterion, ratings of the incomplete information conditions as predictors and the speakers' sex as control variable. We expected these predictors to be intercorrelated and distorted due to multicollinearity. For this reason we also calculated so-called relative weights. Relative weights or relative importance weights range between 0 and 1, are unaffected by multicollinearity, and therefore very helpful when interpreting each predictors' relative contribution in the regression model (Johnson, 2000; Kraha, Turner, Nimon, Zientek, & Henson, 2012; Lorenzo-Seva, Ferrando, & Chico, 2010). Moreover, bivariate correlations between predictor variables provided insight into the interrelations between ratings of the incomplete information conditions.

Previous studies comparing different communication channels found correlations of .35 or higher between nonverbal cues and personality ratings (Borkenau & Liebler, 1992b; Friedman et al., 2006; Koppensteiner, 2013). We used different modalities as predictors in a regression analyses and assumed their effects to add up. Consequently, we expected the multiple regressions we performed to explain at least 20 percent of overall variance. On the basis of such an effect size, an alpha level of .05, a power level of .8, and five predictors in a priori power analyses suggested an optimal sample size of 51 stimuli.

All statistical analyses were carried out in the program R (R Core Team, 2013).

### 3. Results

Descriptive Statistics are presented in Table 1 and internal consistencies between corresponding items of the questionnaire in Table 2. Coefficients ranged from .06 to .94 and thus were unacceptably low in some cases. In particular, the item pairs anxious–easily upset and calm–emotionally stable did not combine properly in the stick figure, in the original voice, nor in the full channel condition. This might have been due to a weakness of the questionnaire on this dimension or to the choice of stimuli. Politicians on a public stage may rarely show body movements or produce vocal cues in which people perceive qualities associated with the items anxious and easily upset. To sum up, internal consistencies we obtained for emotional stability were very low and for this reason we discuss results for this personality dimension on the basis of single items (Tables 3–5). Statistical analyses of the relationships between incomplete information conditions and the full channel conditions are presented as correlation coefficients, relative weights, and regression estimates (see Tables 3 and 4).

Results for extraversion revealed a strong relationship of the ratings in the stick-figure condition and the original voice condition with the ratings in the full channel condition. In addition, there was also a noteworthy relationship between ratings in the static visual condition and the full channel condition. These relationships were reflected in all types of analyses we applied (i.e., see correlations in Table 3, and estimates and relative weights in Table 4). We therefore came to the conclusion that in our setting, judgments of extraversion were mainly affected by the speakers’ voice and by their body motion as well—as to minor degree—by the speakers’ appearances. The content of the speeches had no important effect on ratings of extraversion.

Perceived conscientiousness in the full channel condition yielded high correlation coefficients with ratings of conscientiousness in the static visual condition, in the original voice condition, and the artificial voice condition (i.e., computer voice presenting speech content). Similar results were obtained for estimates of the multiple regression and the relative weights (Tables 3 and 4). Consequently, impressions of conscientiousness were mostly guided by vocal information. However, appearance cues and speech content also played a role. Body motion was revealed as the weakest predictor.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>Pic</th>
<th>Stick</th>
<th>Vox</th>
<th>Covox</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extraversion</td>
<td>10.26</td>
<td>9.65</td>
<td>14.25</td>
<td>15.37</td>
<td>16.28</td>
</tr>
<tr>
<td>(27.84)</td>
<td>(19.89)</td>
<td>(25.66)</td>
<td>(27.76)</td>
<td>(11.71)</td>
<td></td>
</tr>
<tr>
<td>Conscient.</td>
<td>37.77</td>
<td>42.20</td>
<td>29.90</td>
<td>44.79</td>
<td>47.77</td>
</tr>
<tr>
<td>(27.38)</td>
<td>(30.86)</td>
<td>(24.68)</td>
<td>(29.74)</td>
<td>(22.02)</td>
<td></td>
</tr>
<tr>
<td>Openness</td>
<td>−.77</td>
<td>−4.34</td>
<td>−2.68</td>
<td>−.88</td>
<td>3.02</td>
</tr>
<tr>
<td>(14.89)</td>
<td>(20.37)</td>
<td>(11.89)</td>
<td>(15.40)</td>
<td>(13.95)</td>
<td></td>
</tr>
<tr>
<td>Agreeableness</td>
<td>−5.04</td>
<td>−7.29</td>
<td>−3.30</td>
<td>−5.89</td>
<td>−11.59</td>
</tr>
<tr>
<td>(21.68)</td>
<td>(17.99)</td>
<td>(23.03)</td>
<td>(23.48)</td>
<td>(19.92)</td>
<td></td>
</tr>
<tr>
<td>Calm</td>
<td>4.18</td>
<td>3.94</td>
<td>2.58</td>
<td>5.06</td>
<td>7.78</td>
</tr>
<tr>
<td>(21.21)</td>
<td>(16.96)</td>
<td>(19.05)</td>
<td>(22.77)</td>
<td>(16.12)</td>
<td></td>
</tr>
<tr>
<td>Anxious</td>
<td>−18.18</td>
<td>−22.33</td>
<td>−20.61</td>
<td>−27.52</td>
<td>−28.59</td>
</tr>
<tr>
<td>(18.44)</td>
<td>(16.97)</td>
<td>(14.30)</td>
<td>(20.93)</td>
<td>(18.13)</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Values in parenthesis are standard deviations. Full = full channel condition (i.e., all visual and vocal information available), Pic = static visual condition (i.e., single frames of videos), Stick = stick figure condition (i.e., stick-figure animations of body movements), Vox = original voice condition (i.e., voice only), Covox = artificial voice condition (i.e., contents read by computer voice). Conscient. = conscientiousness; emotional stability is represented by the items calm (i.e., calm, emotionally stable) and anxious (i.e., anxious, easily upset). N = 60.
Table 2
Reliability measures of the big five personality dimensions for each of the five experimental conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Extraversion</th>
<th>Conscientiousness</th>
<th>Openness</th>
<th>Agreeableness</th>
<th>Emotional stability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'reserved, quiet' &amp; 'extraverted, enthusiastic'</td>
<td>'disorganized, careless' &amp; 'dependable, self-disciplined'</td>
<td>'conventional, uncreative' &amp; 'open to new experiences, complex'</td>
<td>'critical, quarrelsome' &amp; 'sympathetic, warm'</td>
<td>'anxious, easily upset' &amp; 'calm, emotionally stable'</td>
</tr>
<tr>
<td>Static visual condition</td>
<td>.71</td>
<td>.85</td>
<td>.93</td>
<td>.84</td>
<td>.73</td>
</tr>
<tr>
<td>Stick figure condition</td>
<td>.94</td>
<td>.82</td>
<td>.69</td>
<td>.92</td>
<td>.06</td>
</tr>
<tr>
<td>Original voice condition</td>
<td>.94</td>
<td>.87</td>
<td>.69</td>
<td>.79</td>
<td>.46</td>
</tr>
<tr>
<td>Artificial voice condition</td>
<td>.63</td>
<td>.71</td>
<td>.76</td>
<td>.69</td>
<td>.68</td>
</tr>
<tr>
<td>Full channel condition</td>
<td>.92</td>
<td>.88</td>
<td>.86</td>
<td>.76</td>
<td>.41</td>
</tr>
</tbody>
</table>

Notes. Based on the Spearman–Brown reliability coefficient (i.e., $\rho_{xy} = 2r_{xy}/(1 + r_{xy})$). $N = 60$.

Table 3
Bivariate correlations of the big five ratings in the incomplete information conditions with the big-five ratings in the full channel condition.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Extraversion</th>
<th>Conscient.</th>
<th>Openness</th>
<th>Agreeableness</th>
<th>Emotional stability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Calm, emotionally stable</td>
</tr>
<tr>
<td>Static visual condition</td>
<td>.39***</td>
<td>.51***</td>
<td>.57***</td>
<td>.34***</td>
<td>.30</td>
</tr>
<tr>
<td>Stick figure condition</td>
<td>.67***</td>
<td>.28</td>
<td>.09</td>
<td>.49***</td>
<td>.43***</td>
</tr>
<tr>
<td></td>
<td>[.51, .79]</td>
<td>[.02, .50]</td>
<td>[.17, .33]</td>
<td>[.27, .66]</td>
<td>[.20, .62]</td>
</tr>
<tr>
<td>Original voice condition</td>
<td>.89***</td>
<td>.70***</td>
<td>.66</td>
<td>.89***</td>
<td>.81***</td>
</tr>
<tr>
<td></td>
<td>[.82, .93]</td>
<td>[.54, .81]</td>
<td>[.49, .79]</td>
<td>[.82, .93]</td>
<td>[.69, .88]</td>
</tr>
<tr>
<td>Artificial voice condition</td>
<td>.24</td>
<td>.54***</td>
<td>.28</td>
<td>.50**</td>
<td>.48***</td>
</tr>
<tr>
<td></td>
<td>[.01, .47]</td>
<td>[.33, .70]</td>
<td>[.03, .50]</td>
<td>[.28, .67]</td>
<td>[.26, .65]</td>
</tr>
</tbody>
</table>

Notes. Numbers in brackets are 95% confidence intervals. Artificial voice condition is speech content read by a computer voice. Results for emotional stability are presented on the level of single items because of low reliability (see Table 1). Conscient. = conscientiousness. $N = 60$.

* $p < .05$.
** $p < .005$.
*** $p < .001$.

Table 4
Multiple regression analyses using the big five ratings in full channel condition as criterion and big five ratings of the incomplete information conditions as predictors.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Extraversion</th>
<th>Conscient.</th>
<th>Openness</th>
<th>Agreeableness</th>
<th>Emotional stability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Calm, emotionally stable</td>
</tr>
<tr>
<td>Sex</td>
<td>- .22 (.03)</td>
<td>-1.88 (4.37)</td>
<td>.50 (2.87)</td>
<td>-3.42 (2.39)</td>
<td>3.06 (2.00)</td>
</tr>
<tr>
<td></td>
<td>[-6.44, 5.49]</td>
<td>[-10.08, 6.93]</td>
<td>[-5.86, 5.23]</td>
<td>[-7.95, 1.00]</td>
<td>[-2.72, 8.84]</td>
</tr>
<tr>
<td></td>
<td>.02</td>
<td>.00</td>
<td>.05</td>
<td>.06</td>
<td>.01</td>
</tr>
<tr>
<td>Static visual condition</td>
<td>.12 (.08)</td>
<td>.28*** (.07)</td>
<td>.32*** (.07)</td>
<td>.23*** (.07)</td>
<td>.20 (.09)</td>
</tr>
<tr>
<td></td>
<td>[.03, -.28]</td>
<td>[.14, .42]</td>
<td>[.21, .45]</td>
<td>[.11, .34]</td>
<td>[.04, .38]</td>
</tr>
<tr>
<td></td>
<td>.06</td>
<td>.16</td>
<td>.23</td>
<td>.07</td>
<td>.06</td>
</tr>
<tr>
<td>Stick figure condition</td>
<td>.24</td>
<td>.03</td>
<td>.00</td>
<td>.19</td>
<td>.11</td>
</tr>
<tr>
<td>Original voice condition</td>
<td>.72*** (.07)</td>
<td>.49*** (.08)</td>
<td>.54*** (.10)</td>
<td>.74*** (.07)</td>
<td>.62*** (.08)</td>
</tr>
<tr>
<td></td>
<td>[.56, .85]</td>
<td>[.32, .65]</td>
<td>[.34, .71]</td>
<td>[.61, .88]</td>
<td>[.46, .76]</td>
</tr>
<tr>
<td></td>
<td>.50</td>
<td>.34</td>
<td>.31</td>
<td>.52</td>
<td>.47</td>
</tr>
<tr>
<td>Artificial voice condition</td>
<td>.02 (.13)</td>
<td>.30*** (.11)</td>
<td>.01 (.11)</td>
<td>.00 (.07)</td>
<td>.14 (.11)</td>
</tr>
<tr>
<td></td>
<td>.12</td>
<td>.14</td>
<td>.04</td>
<td>.11</td>
<td>.11</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.84</td>
<td>.67</td>
<td>.63</td>
<td>.86</td>
<td>.76</td>
</tr>
</tbody>
</table>

Notes. Numbers are regression estimates with SEs in parenthesis. Numbers in brackets are 95% confidence intervals (based on 9999 bootstrapped replicates). Bold numbers are relative weights, which provide explained variance of single predictor without influence of other predictors. Artificial voice condition is speech content read by a computer voice. Results for emotional stability are presented on the level of single items because of low reliability (see Table 1). Conscient. = conscientiousness. $N = 60$.

* $p < .05$.
** $p < .005$.
*** $p < .001$. 
Ratings of perceived openness in the full channel condition were strongly related to ratings in the original voice condition and the static visual condition (Tables 3 and 4). Correlation coefficients also provided a noteworthy effect size for speech content (Table 3). In conclusion, participants mainly tended to ascribe openness to the speaker’s appearances and their voices.

Full channel ratings of agreeableness showed strong bivariate correlations with ratings in the original voice condition, ratings in the stick-figure condition, ratings in the artificial voice condition and a noteworthy one for ratings in the static visual condition (see Table 3). These patterns of relationships were not in accordance with the β-weights of the multiple regression (Table 4). It only provided strong estimates for the static visual condition and the original voice condition. This was an indicator of multicollinearity, which distorts β-weights and, therefore, undermines accurate interpretation of the data. Relative weights, which are unaffected by multicollinearity, replicated the patterns found by the bivariate correlations (Table 3). Thus, we concluded that ratings of agreeableness were guided by appearance cues, by body motion, by vocal cues and by the verbal content the speakers presented.

Due to the low internal consistencies results for emotional stability were analyzed on the level of single items. The item pair calm—emotionally stable yielded notable correlation coefficients and relative weights for all conditions, whereas regression estimates labeled ratings of the original voice condition as strong predictors and ratings of body motion and appearance as predictors with a moderate influence (Tables 3 and 4). The item pair anxious—easily upset provided a strong correlation of full channel estimations (Table 3). This was mostly due to paralinguistic cues such as intonation and pitch because the strong impact of the speakers’ voice on personality judgments did not disappear when speech content was controlled. Other studies also revealed that vocal cues are an important influential factor in impression formation (Borkenau & Liebler, 1992a; Friedman et al., 2006). However, they did not play such a predominant role as in our study. Participants were, of course, aware that they were judging politicians. For this reason, they might have attended more to the speaker’s voices than to other information. The majority of studies that compared the impact of different communication channels did not use politicians for their experiments or instructed actors to display certain behaviors (e.g., Awamleh & Gardner, 1999; Borkenau & Liebler, 1992a; Friedman et al., 2006). This is different from our study design and might explain why vocal cues were a strong contributor. Actors may overdo their acting and display stereotypical behaviors thereby drawing attention to cues that are less salient when real politicians present themselves. On the other hand, judging politicians may raise different expectations than judging people drawn from an average population.

4. Discussion

4.1. Original voice condition

The most prominent finding of our study was that all personality ratings of the original voice condition, in which the participants only listened to an audio track of the selected speeches, were strongly related to corresponding ratings of the full channel condition, in which participants watched the unaltered original video clips. This was mostly due to paralinguistic cues such as intonation and pitch because the strong impact of the speakers’ voice on personality judgments did not disappear when speech content was controlled. Other studies also revealed that vocal cues are an important influential factor in impression formation (Borkenau & Liebler, 1992a; Friedman et al., 2006). However, they did not play such a predominant role as in our study. Participants were, of course, aware that they were judging politicians. For this reason, they might have attended more to the speaker’s voices than to other information. The majority of studies that compared the impact of different communication channels did not use politicians for their experiments or instructed actors to display certain behaviors (e.g., Awamleh & Gardner, 1999; Borkenau & Liebler, 1992a; Friedman et al., 2006). This is different from our study design and might explain why vocal cues were a strong contributor. Actors may overdo their acting and display stereotypical behaviors thereby drawing attention to cues that are less salient when real politicians present themselves. On the other hand, judging politicians may raise different expectations than judging people drawn from an average population.
4.2. Speech content

Judgments in the original voice condition did not only convey speech content but also information about intonation, voice pitch, and other qualities of the voice. To separate such vocal cues from speech content, we also collected ratings of the speeches' wording read by a computer voice and related these ratings to ratings of full channel condition. As expected full channel judgments of conscientiousness, but also judgments of agreeableness and one item pair of the personality dimension emotional stability (i.e., calm, emotionally stable) could be predicted by speech content to a certain degree. In other words, participants read some traits and qualities into what the speakers said and not only how they said it. Unlike some other studies we were unable to show that verbal content dominates over nonverbal cues in judgments of politicians (Krauss et al., 1981; Nagel et al., 2012). Nevertheless, the results obtained indicate that people integrate speech content when forming first impressions.

4.3. Static visual information

Static visual information, which was presented as single pictures, yielded notable relationships with the full channel condition. Previous research has already shown that appearance cues alone—clothing styles, physiognomic features and facial expressions, just to give a few examples—affect people's first impressions (e.g., Naumann et al., 2009). We have supported such findings by demonstrating that static visual cues contribute to overall impressions of a person. Consequently, people form an impression on the basis of static cues before any motion behavior occurs or a word is spoken. The documented effects were particularly strong for ratings of conscientiousness and openness. Participants had no information about party membership, but static visual cues may already reflect a conservative or a more progressive mindset and this may more strongly guide attributions of openness and conscientiousness than attributions of other personality dimensions. Future studies including facial expressions, physiognomic features, and clothing style as variables in the analyses could support such assumptions.

4.4. Body motion

The study's main focus was to estimate the relative role of body motion when people judge politicians giving a speech. To accomplish this, we translated the speakers' body motion into animated stick-figures. Full channel ratings of extraversion, agreeableness, and an item of emotional stability were strongly related to corresponding ratings of these abstract stimuli. Although other communication channels, in part, showed stronger effects over a wider range of ratings, it is still surprising that a parsimonious cue, as simple as a stick-figure, can guide social perceptions. In previous research, where we applied the same method of motion capturing, we revealed that people associate personality traits with salient behaviors (Krauss et al., 1981; Nagel et al., 2012). Nevertheless, the results obtained indicate that people integrate speech content when forming first impressions.

5. General discussion

We randomly extracted short video clips from speeches that were given in the German parliament. These clips were then turned into five different types of stimuli, which were rated on the Big Five personality taxonomy. Stimuli were either (1) short extracts of the original videos, (2) audio sequences of the speeches, (3) speaker's body movements turned into stick-figures, (4) single pictures of the speakers, or (5) the content of the speeches read by a computer voice. To estimate the influence of these different communication channels on impression formation we related ratings in the incomplete information conditions to ratings in the original full channel video clip condition.

5.1. The role of different communication channels

People make sense of brief displays of nonverbal and verbal information and although such snap judgments can be a source of false beliefs and prejudices, they often convey socially relevant information (e.g., Ambady et al., 2000). Studies examining the influence of different communication channels on impression formation found that people read information into and from a variety of cues. Most research suggests that visual cues have a prevailing influence on social judgments and that speech content only plays a minor role (e.g., Awamleh & Gardner, 1999; Mehrabian, 1972). Other researchers were unable to replicate such findings, which supports the idea that the influence of nonverbal and verbal cues varies according to context and audience involvement (e.g., O'Sullivan et al., 1985; Petty et al., 1983).

In this study, we found a prevailing effect of vocal cues on impression formation but the relative role of different communication channels may, of course, depend on different factors that can hardly be tested and controlled in a single study. Future studies should thus extend the current experimental set-up by...
manipulating the salience of cues (e.g., manipulating vocal parameters), refine analyses by describing behaviors and appearance cues in more detail (e.g., clothing style), and include different contextual information (e.g., party affiliation, status). Also, politicians acting in the public arena may trigger differential judgments than the stimuli some other researchers used to investigate the impact of different communication channels (e.g., Borkenauf & Liebler, 1992a; Friedman et al., 2006). As in our study, Krauss et al. (1981) as well as Nagel et al. (2012) collected ratings of politicians and found that people mainly judge them by verbal content. This is partly in accordance with our findings because we also revealed a non-negligible relationship between speech content and some trait ratings. However, the outcomes of the regression analyses suggest that, in our study, prosody and voice qualities had a marked impact on the participants’ judgments.

Overall, the results obtained show that people's first impressions are guided by cues from different communication channels. Put simply, in real life encounters there is no single cue that creates a first impression. Our analyses supports this by showing that in most cases single modalities (e.g., body motion, voice only) explained less variance than combinations of these modalities. Consequently, when making their judgments people seem to rely on a great variety of nonverbal and verbal cues. Also, some modalities such as voice and body motion are intertwined and communicate redundant information to a certain extent, while other modalities do not show such a redundancy. For instance, conscientiousness was more easily attributed to speech content than extraversion. Irrespective of that our results support the idea that people’s first impressions are guided by information from different communication channels (Bänziger et al., 2012; Mehu & van der Maaten, 2014).

5.2. The relative role of body motion

By translating the speakers’ nonverbal behavior into animated stick-figures, we created an abstract stimulus that served as representative of body motion. It has already been shown that parsimonious cues displaying motion behavior communicate socially relevant information (e.g., Blake & Shiffrar, 2007; Chouchourelou et al., 2006). In this study, we extended such research by comparing the influence of cues conveyed by a moving body with information from other communication channels. We were able to show that, for perceptions of extraversion and agreeableness, body motion is a major player. Given that our stick-figure animations reduce information to an abstract stimulus with an artificial appearance, the results obtained are impressive. They indicate people do not only ascribe intentions and personality traits to abstract representatives of body motion, but they also use motion cues to form impressions of their interaction partners.

Rapidly categorizing another individual’s intentions, traits and behavioral tendencies may be an ability that has been formed during human evolution (Buss & Greiling, 1999; Fiske, Cuddy, & Glick, 2007). However, some traits are assumed to be of higher social relevance in first encounters because they inform the decision whether to approach or avoid an individual. For instance, Oosterhof and Todorov (2008) found that judgments of neutral faces are reducible to the independent dimensions of valence (i.e., represented by facial trustworthiness) and dominance. They further suggest that evaluations on these dimensions are overgeneralizations for inferring harmful intentions. Other research has shown that attributions of extraversion to a stick-figure animation show a strong relationship with perceived dominance, while perceived agreeableness is strongly related to perceived trustworthiness (Koppensteiner, Stephan, & Jäschke, 2015). This supports the conclusion that people have a higher sensitivity for cues of extraversion and agreeableness because these belong to basic categories of social evaluation with a direct link to survival. For this reason, overgeneralizations with regard to potential threats may not only occur when faces are assessed but also when people categorize body motion. Furthermore, in contrast to facial information, body motion is recognizable from greater distances. It is thus conceivable that motion cues are socially relevant because they allow decision making as another individuals approach, before these individuals come too close.

5.3. Future directions

In real life situations, nonverbal and verbal information does not disintegrate into pieces. Our results, and other research, suggest that some cues are intertwined. A well-known link, for instance, is between speech and gesturing (e.g., McNeill, 1985). Gestures are sometimes used to emphasize what is being said or go together with voice intonation. Hence, an overlap in the communicative value of motion cues, speech content, and voice quality is well established and the results obtained in the present study provide further support for this. Future research could investigate such an overlap in more detail by extracting certain motion cues as we have done in previous studies (see Koppensteiner & Grammer, 2010) and relating them to vocal cues. Follow-up studies with a more in depth analyses on the interrelations between different cues would also build a bridge to research that investigates how observers react to violations of such linkages (Weisbuch, Ambady, Clarke, Achor, & Weelee, 2010). Presenting contradictory nonverbal and verbal information could show that there is indeed redundancy and that different cues might form one single cue with a common communicative value. This might also provide further support for work showing that judgments on the basis of multimodal information are more accurate (Bänziger et al., 2012; Mehu & van der Maaten, 2014).

A next step in this kind of research would be to investigate how different nonverbal information influences decision making on important issues. Researchers have already shown that facial photographs from political candidates can be used to predict hypothetical or actual election outcomes (e.g., Antonakis & Dalgas, 2009; Little et al., 2012; Olivola & Todorov, 2010). Moreover, other studies reveal a link between vocal cues and leadership qualities (Klofstad et al., 2012; Tigue et al., 2012). Taking these as a starting point future research could extend into the domain of leadership research to examine how the interplay of different cues or multimodal cues affects decision-making.

5.4. Conclusions

In summary, our findings indicate that people make use of verbal and different nonverbal cues when forming first impressions of politicians. Static visual cues, body motion, verbal content, and vocal cues were related to people’s impressions. Vocal cues were revealed as a strong predictor for all trait ratings in our study, yet other communication channels also proved important. The impact of body motion was particularly strong for ratings of agreeableness and extraversion were also linked to corresponding ratings based on the speakers’ voices. In other words, with regard to these personality dimensions body motion is coupled to information from the vocal channel. Cues communicating agreeableness and extraversion may reflect social abilities, which are helpful in social encounters and facilitate establishing interactions. Hence, people may be more sensitive to cues of extraversion and agreeableness; for this reason, they also read these traits into body motion.


