Running head: DYNAMIC VOCAL SIGNALS PREDICT SOCIAL RANK

Listen, Follow Me:

Dynamic Vocal Signals of Dominance Predict Emergent Social Rank in Humans

Word Count: 7,571

Abstract

Similar to the nonverbal signals shown by many non-human animals during aggressive conflicts, humans display a broad range of behavioral signals to advertise and augment their apparent size, strength, and fighting prowess when competing for social dominance. Favored by natural selection, these signals communicate the displayer's capacity and willingness to inflict harm, and increase responders' likelihood of detecting and establishing a rank asymmetry, and thus avoiding costly physical conflicts. Included among this suite of adaptations are vocal changes, which occur in a wide range of non-human animals (e.g., chimpanzees, rhesus monkeys) prior to aggression, but have not been systematically examined in humans. The present research tests whether and how humans use vocal pitch modulations to communicate information about their intention to dominate or submit. Results from Study 1 demonstrate that in the context of face-to-face group interactions, individuals spontaneously alter their vocal pitch in a manner consistent with rank signaling. Raising one's pitch early in the course of an interaction predicted lower emergent rank, whereas deepening one's pitch predicted higher emergent rank. Results from Study 2 provide causal evidence that these vocal shifts influence perceptions of rank and formidability. Together, findings suggest that humans use transient vocal changes to track, signal, and coordinate status relationships.

Subject Areas:

behavior, evolution

Keywords:

dominance, social hierarchy, vocal pitch, nonverbal behavior, signaling

Although coordination and affiliation form the primary fabric of social life among many species, group living necessarily results in conflicts over divergent goals, and zero-sum competitions for valued resources. To tackle this recurrent problem, many species have evolved psychological and behavioral adaptations that facilitate the formation and maintenance of rank asymmetries (Schjelderup-Ebbe, 1922; Lorenz, 1966; Mazur, 1985; Brown, 1991). The resultant status hierarchies establish a mutually accepted agreement on the differential priority and access to contested resources, and thereby enable stable patterns of social exchange, prevent costly fights, and, in consequence, maximize individual fitness (Christian, 1970; Maynard Smith & Parker, 1976; Maynard Smith & Price, 1973; Parker, 1974). Indeed, a large body of evidence indicates that hierarchical stratification organizes many everyday social exchanges and facilitates group coordination. In humans, relative social rank reliably predicts patterns of dominance-submission (i.e., influence, deference, attention; Chance, 1967; Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013; Fournier, Moskowitz, & Zuroff, 2002; Thomsen, Frankenhuis, Ingold-Smith, & Carey, 2011; Tiedens & Fragale, 2003), distribution of wealth (Hawley, 2002; Sidanius & Pratto, 1999), health (Adler et al., 1994), access to mates (Buss, 1989), and reproductive success (von Rueden, Gurven, & Kaplan, 2011). Equally important, stable hierarchical systems enhance group-wide motivation, cooperation, and productivity (Anicich, Swaab, & Galinsky, 2015; Bendersky & Hays, 2012; Halevy, Chou, Galinsky, & Murnighan, 2012; Ronay, Greenaway, Anicich, & Galinsky, 2012; Tiedens & Fragale, 2003).

The substantial advantages conferred by social rank asymmetries—combined with evidence of spontaneous, rapid, and reliable emergence of hierarchical relationships within human social groups across a broad range of environments (Bass, 1954; Kalma, 1991; Lee, 1979; Lewis, 1974)—leads to the expectation that our species be equipped with psychological and behavioral adaptations designed to facilitate the signaling and detection of relative rank differences. In fact, several externally visible features of human morphology, including height, muscularity, and facial structure, reliably track and signal fighting ability, and are used to assess formidability (Carré & McCormick, 2008; Fessler, Holbrook, & Snyder, 2012; Judge & Cable, 2004; Puts, Gaulin, & Verdolini, 2006; Sell et al., 2009, 2010).

However, these stable features have several limitations. For example, when disparities in size and strength are not readily apparent (Parker, 1974), or contestants' ability to inflict harm is heavily influenced by ecology-specific factors—including availability of allies, existing injuries, differences in age and overall physiological condition, territorial ownership and experience, or (in humans) differential access to weapons, and disparate wealth and resources that correlate with fighting ability—these relatively immutable morphological features become inadequate for signaling an individual's current competitive intentions. As a result, natural selection should also favor the emergence of more dynamic behavioral displays to flexibly advertise size-independent rank-attainment motivations; these displays, in conjunction with stable morphological features, would provide a summary assessment of individuals' relative formidability under current ecological conditions (Bro-Jørgensen, 2010; Otte, 1974; Tinbergen, 1959). Consistent with this reasoning, as well as observations made of other primates (de Waal, 1982; de Waal & Luttrell, 1985), a growing body of evidence indicates that humans rely on dynamic momentary behavioral displays to communicate and exchange information about formidability and competitive or submissive intentions. Among the most visually detectible are postural expansion versus constriction (Dovidio et al., 1988; Shariff & Tracy, 2009; Tiedens & Fragale, 2003), emotion expressions of pride, anger, and contempt versus fear, shame, and sadness (Fournier, 2009; Keltner, Gruenfeld, & Anderson, 2003; Marsh, Cardinale, Chentsova-Dutton, Grossman, &

Krumpos, 2014; Shariff, Tracy, & Markusoff, 2012; Tracy, Shariff, Zhao, & Henrich, 2013), and sustained eye gaze versus gaze aversion (Dovidio et al., 1988).

Despite the importance of these visually detectible signals, there are occasions when even these more dynamic cues would be ineffective, due to distance, darkness, or obscuration (Sell et al., 2010). These limitations may have favored the emergence of supplemental signaling systems that do not rely on sight, such as vocal signals. Indeed, several lines of research indicate that individual differences in habitual vocal pitch—the perceptual parameter that corresponds to vocal fundamental frequency, and the most perceptually salient aspect of the human voice (Banse & Scherer, 1996)—are consistently linked to rank-attainment motivations and outcomes in humans. First, deeper voices convey a speaker's increased threat potential, by virtue of their reliable association with stable characteristics related to threat, including physical size and upper-body strength (Bruckert, Liénard, Lacroix, Kreutzer, & Leboucher, 2006; Evans, Neave, & Wakelin, 2006; Hodges-Simeon, Gurven, Puts, & Gaulin, 2014; Puts, Apicella, & Cárdenas, 2012), endogenous circulating testosterone (Bruckert et al., 2006; Dabbs Jr. & Mallinger, 1999; Evans, Neave, Wakelin, & Hamilton, 2008; Puts et al., 2012) and exposure to testosterone during development (Harries, Hawkins, Hacking, & Hughes, 1998). Studies have found that increases in testosterone, such as those resulting from experimental administrations, leads to heightened status-seeking motivations, increased social vigilance, and reductions in fear and stress (see Eisenegger, Haushofer, & Fehr, 2011). Thus, a deepening voice may cue that an individual is psychologically and physiologically ready for a status competition.

Second, experimental studies have found that lower pitched voices are widely perceived as indicating greater size, strength, dominance, and leadership capacity (Feinberg, Jones, Little, Burt, & Perrett, 2005; Gregory Jr., Dagan, & Webster, 1997; Gregory Jr., Green, Carrothers, Dagan, & Webster, 2001; Klofstad, Anderson, & Peters, 2012; Puts et al., 2006; Puts, Hodges, Cárdenas, & Gaulin, 2007; Tigue, Borak, O'Connor, Schandl, & Feinberg, 2012; Wolff & Puts, 2010); but see Ko, Sadler, & Galinsky, 2015). These findings, along with evidence of a similar association between the pitch of acoustic signals and fighting ability in many other species (Gerhardt, 1994), suggest that individual differences in habitual vocal pitch may provide cues about the stable fighting ability or propensity of the vocalist.

Dynamic *changes* in vocal pitch, however, have received far less empirical attention, compared to stable individual differences in habitual pitch frequency (Puts et al., 2006; Scherer, 1986). Comparative studies, however, have documented the pervasiveness of dynamically altered acoustic signals across the animal kingdom. Frogs, for example, lower the pitch of their calls during aggressive encounters, particularly under environmental conditions where it would be especially advantageous to exaggerate one's size (e.g., to deter intruders; Bee, Perrill, & Owen, 1999, 2000), and these deepening pitch calls are more effective at deterring challengers than calls that either increase or do not change in pitch frequency (Wagner Jr., 1992). Similarly, both rhesus monkeys and chimpanzees produce acoustically distinct vocalizations depending on the relative rank of the opponent with whom they are fighting and the ally whose help they are seeking (Gouzoules, Gouzoules, & Marler, 1984; Morton, 1977; Slocombe & Zuberbühler, 2007).

These patterns found in the natural world converge with one of few relevant studies examining pitch change in humans, which found that men use may systematically use pitch modulations to signal dominance. Specifically, Puts and colleagues (2006) found that, when addressing an ostensible rival, men who considered themselves to be physically stronger lowered their vocal pitch, whereas men who considered themselves to be physically weaker raised their pitch. This finding is consistent with results from several studies examining pitch modulations in naturalistic interactions, which demonstrate that lower-status counterparts of an interaction show asymmetrically greater variation in pitch, in an effort to accommodate a person of higher status (Gregory Jr. et al., 2001; Gregory Jr. & Webster, 1996). This research, along with studies linking stable individual differences in vocal pitch to social rank in humans, raise the possibility that similar pitch alteration signals may function to effectively communicate the intention to dominate or submit in our species. If this is the case, systematic patterns of pitch alteration that occur during human rank contests may influence the outcome of these disputes and shape subsequent rank asymmetries. Although these studies tentatively suggest the presence of systematic pitch alterations in human rank competitions, it remains to be determined whether and how these shifts in pitch influence perceptions of the signaler's dominance motivations, and whether they shape resultant rank asymmetries.

To examine this issue, we tested whether people spontaneously show dynamic changes in their vocal pitch during a rank contest, and how these emitted cues shape the subsequently emerging social hierarchy. We predicted that a dynamically deepening pitch profile communicates an individual's motivation to dominate, and, as a result, would be associated with higher emergent rank. Conversely, we expected a rising pitch profile to communicate the willingness to submit, and thus be associated with lower emergent rank. Furthermore, because pitch modulations are expected to shape rank outcomes primarily by augmenting or diminishing one's apparent physical size, motivation to engage in conflict, and formidability—given the widespread tendency to ascribe greater size and strength to deeper voices (Feinberg, Jones, Little, et al., 2005; Puts et al., 2006)—pitch modulation should be associated with perceptions of *dominance*, in particular, a rank-attainment strategy based on threat of force and intimidation

(Cheng et al., 2013; Cheng, Tracy, & Henrich, 2010a; Henrich & Gil-White, 2001). We further expected that the rank consequences of pitch alterations would *not* be attributable to prestige—an alternative rank attainment strategy based on earned respect and the demonstration of skills or expertise. In light of experimental evidence indicating that the perceptual association between deeper vocal pitch and dominance applies to both male and female voices (Borkowska & Pawlowski, 2011; Jones, Feinberg, DeBruine, Little, & Vukovic, 2010), we expected these predictions to hold across gender (notably, testosterone surges have been shown to have similar effects in men and women).

We tested these predictions in two studies. Study 1 sought to establish the presence of rank-related pitch alterations in a real-world context by assessing naturally occurring vocal changes during the process of hierarchy formation in small groups. Study 2 complemented this ecologically valid approach by manipulating pitch alterations via a playback paradigm, and testing whether these shifts influenced listeners' perceptions of rank-seeking intentions and dominance.

Study 1: Pitch Alterations and Rank Attainment in Face-to-Face Social Interactions

Study 1 tracked naturally occurring dynamic changes in pitch during face-to-face interactions among individuals assigned to work together on a problem-solving task. We tested whether vocal changes that spontaneously occurred during the first moments of these interactions predicted individuals' likelihood of prevailing or submitting in the group, as well as perceptions of formidability.

Method

Participants. One hundred, ninety-one students (54% male), ranging in age from 17 to 52 (*M*=23.01, *SD*=6.08), at a large Canadian university participated for monetary compensation.

Participants were randomly assigned to 1 of 36 same-sex groups that ranged in size from 4 to 7 individuals (*M*=5.34, *SD*=.83). No group members were previously acquainted.

Procedure. The study began with participants completing a decision-making task privately. They then worked collectively with their group for 20 minutes on the same task. They were instructed to use their previously completed private responses to guide these group discussions. The task, known as "Lost on the Moon" (Bottger, 1984), involves rank-ordering 15 items (e.g., oxygen tanks, heating unit) according to their perceived utility for surviving a disaster on the moon. The group component was video-recorded. Participants' task engagement was incentivized via the instruction that every member of high-performing groups would receive a $5 bonus, in addition to their $10 compensation for participation. In actuality, all participants received this monetary bonus. After the group task, participants privately rated the social rank and formidability, a trait central to dominance (Henrich & Gil-White, 2001), of each group member.

Measures.

Vocal pitch**.** An audio track was created for each group session from the video-recording, and saved as an uncompressed 'wave' file with 44.1 kHz sampling rate and 16-bit quantization. We examined participants' spontaneous speech in the initial minutes of the group task by sampling each participant's first three utterances, which were spoken, on average, at 3.28, 4.32, and 5.98 minutes into the 20-minute collaborative component. We focused on these initial utterances to track changes in pitch that occurred largely before rank asymmetries were established. Furthermore, prior work suggests that dominance signals are likely to be especially pronounced and readily attended to early on in a social interaction, before asymmetries are determined (Curhan & Pentland, 2007; Mazur, 1985).

An utterance was defined as uninterrupted speech lasting 40-ms or longer. This short time frame was adopted because, in these naturalistic interactions, many participants made only brief comments (e.g., “okay”, “sure”), particularly at the beginning of the task. The fundamental frequency of each utterance was analyzed with Praat phonetic analysis software.

Social rank**.** Emergent social rank was measured using three indices: group member-rated rank, outside observer-rated rank, and behavioral decision-making impact.

(1) Group member-rated social rank was derived from peers’ post-task ratings of every other group member, on three items: “this person led the task,” “this person had high status,” and “this person was paid attention”. Ratings were made on a 7-point scale (1=*Not at all*; 7=*Very much*). We used the Social Relations Model (Kenny, 1994) to calculate the *target effect* for each item, to capture each target’s average rating after statistically removing idiosyncratic perceiver and dyadic relationship biases. The target effect scores of these items were averaged to form an overall measure of group-member rated rank (inter-item α=.89).

(2) Outside observer-rated rank was derived from judgments made by female undergraduate research assistants who were blind to our hypotheses. They watched the video-recordings of the group interactions in isolation and independently responded to the question “how influential was this individual over other group members,” on a 5-point scale (1=*Not at all*; 5=*Extremely*; inter-rater α=.87). Due to the large number of group recordings, two observers coded one half, and two other observers coded the other half. Ratings provided by the two observers for each participant were averaged to create an index of outside observer-rated rank.

(3) Behavioral decision-making impact was quantified as the degree to which each participant steered the group’s collective decision closer to his/her own (Cartwright, 1959), by computing an index of similarity between each participant’s initial *private* response on the task

and the group's final *collective* response [see online supplementary material (OSM) for details; Bottger, 1984].

***Dominance and Prestige*.** After completing the group-based decision-making task, participants rated each of the other group members on previously validated scales designed to assess perceptions of dominance and prestige (Cheng, Tracy, & Henrich, 2010b). These scales assess individuals' fear (e.g., "I'm afraid of her") and respect (e.g., "I respect and admire her") toward others, and directly tap the interpersonal perceptions that indicate the effective establishment of a formidable or admirable appearance. Ratings were made on a 7-point Likert scale. We again used the Social Relations Model (Kenny, 1994) to compute target effect scores for each item, and then aggregated across the 8 dominance items (inter-item α=.93) and the 8 prestige items (inter-item α=.89) to form an overall measure of dominance (M=2.33; SD=.81) and prestige (M=4.93; SD=.62).

Table S2 provides descriptive information on all key variables.

Results and Discussion

Pitch Modulation and Emergent Social Rank. To obtain a composite measure of emergent social rank for each participant, we computed standardized values (with mean of zero and standard deviation of one) for each of the three indices of social rank, then averaged across them (α=.68). We focus on this rank composite in our primary analyses to minimize stochastic variation introduced by methodological differences between the three measures, and to derive a more precise indicator of individuals' social standing that incorporates different aspects of rank. However, the same qualitative results were obtained when we examined each of the three rank indices separately (see OSM).

We analyzed these data using hierarchical linear modeling (HLM) to estimate an individual growth model (Bryk & Raudenbush, 1987). In contrast to OLS models, which capture only fixed effects and assume a single estimate for intercepts and slopes across the entire sample, this analytic approach incorporates random effects and allows for person-specific initial pitch and pitch trajectories. It allows us to examine how vocal pitch, measured across multiple instances at the within-person level, changes over time in the form of a pitch trajectory, and, more importantly, how this trajectory varies as a function of social rank (i.e., the moderator), measured at the between-person level. Although our primary interest was in the effect of initial pitch alterations on subsequent rank, which was addressed by examining the association between pitch alterations in the first few minutes of the interaction and emergent social rank assessed after the end of interaction, rank was treated as a moderator of the relation between time (the predictor) and vocal pitch (the outcome), to model individual growth in the form of pitch trajectories (Bryk & Raudenbush, 1987).

Specifically, we estimated the fundamental frequency of each person's three utterances (treated as the dependent variable in this analysis) as a function of the order of the utterance (hereafter, time; coded from 0, for the first utterance, to 2, for the third utterance), his/her social rank composite score (grand-mean centered), and the cross-level interaction between time and social rank, controlling for gender (with female=1) and its interaction with time. This resulted in a growth model with random intercepts and random slopes to represent each individual's person-specific initial pitch and pitch trajectory, respectively, independent of the effects of gender. We tested the main hypothesis by examining the coefficient on the time × social rank interaction term, which indicates whether and how the pitch trajectories (represented by random slopes in the model) varied as a function of emergent social rank, independent of controls. A non-zero negative interaction would be consistent with the prediction that individuals who lowered their pitch across the initial moments of the interaction eventually acquired high rank, whereas those who raised their pitch during those moments emerged as lower ranking. These analyses include data from 173 participants for whom pitch estimates were available (see OSM for details on missing data).

Results indicated a significant negative time × social rank interaction predicting vocal pitch (b=-4.45, SE=1.33, p=.001, .95CI[-7.05, -1.85]; see Table 1). Although the pitch trajectory of the average participant with moderate social rank showed a non-significant slight positive incline, individuals who emerged as low in rank showed a heightening pitch profile over time, whereas those who emerged as high rank showed a deepening pitch profile. This implies that the pitch trajectory of low ranking men and women (i.e., those whose social rank scores were 1-point below the grand mean) increased respectively by 5.31 Hz and 7.79 Hz over each subsequent utterance, or by 10.62 Hz and 15.58 Hz over the three utterances assessed. In contrast, the pitch trajectory of high ranking men and women (those whose social rank scores were 1-point above the grand mean) *dropped* respectively by 3.59 Hz and 1.11 Hz over each subsequent utterance, or by 7.18 Hz and 2.22Hz over the three utterances assessed. These average magnitudes of pitch increases and decreases exceed the just-noticeable difference threshold for human voices, which psychoacoustic studies have generally revealed to be roughly 2-4 Hz in this frequency range (Ladefoged, 1996; Sinnott, Owren, & Petersen, 1987; Smith, Patterson, Turner, Kawahara, & Irino, 2005). This analysis also indicates that, on average, individuals who eventually acquired higher rank began the interaction with a higher pitch than those who eventually occupied a lower rank; a 1-point increase in social rank was associated with a 6 Hz higher initial pitch (in the first utterance). However, mean pitch parameters (e.g., aggregated pitch across utterances) had little to no predictive power on social rank (see OSM). Mean pitch trajectories also did not differ significantly by gender.

To illustrate this effect, Figure 1 plots the pitch trajectories for individuals who attained high rank (at the 90th percentile) and those who attained low rank (at the 10th percentile) for men and women separately, to capture differences in pitch elevation between the genders. On average, as rank increased from the 10th percentile to the 90th percentile, pitch trajectory changed by approximately -9 Hz per utterance or -18 Hz over three utterances.

These results were robust to a variety of checks, which are presented in detail in the OSM. First, alternative model specifications indicate that the effect is robust to the inclusion of additional controls, including age, height, weight, group size, and postural expansiveness (coded from video-recordings; Table S3). Of these models, the one controlling for gender showed the best fit adjusted for model size (BIC), and is displayed in Table 1 and Figure 1; none of the other controls were significant. Second, reconfirming the predictive importance of the interaction between time and social rank reported above, goodness of fit measures (both AIC and BIC) indicate that across all models examined—with or without controls—the inclusion of the interaction term yielded a better fit (Tables S3 and S4). Third, as noted, the same qualitative results were obtained in alternative models using each of the three separate rank indices, and when controls were included (Table S5). Fourth, additional analyses using the simple difference between the pitch of the third utterance and first utterance (omitting the second)— a potentially more intuitive approach to assessing change (Rogosa & Willett, 1983)—converge with results from the growth model (Table S7).

Pitch Modulation and Perceived Dominance. We predicted that the impact of pitch alterations on rank outcomes occurs by virtue of diminishing or amplifying the vocalizer's perceived size, threat, competitive motivation, and formidability. We therefore predicted that pitch changes would be associated with the signaler's perceived dominance, but not prestige.

To test this account, we ran a separate growth model in which pitch was estimated as a function of time of utterance, peer ratings of dominance (grand-mean centered), peer ratings of prestige (grand-mean centered), and the cross-level interaction between time and both perceptions, controlling for gender and its interaction with time. Table S8 presents detailed results from this baseline model and other models with additional or fewer controls. As predicted, the time × dominance interaction term significantly and negatively predicted vocal pitch, independent of controls (b=-2.84, SE=1.17, p=.016, .95CI[-5.14, -.54]), whereas the time × prestige term lacked significant predictive power (b =-1.07, SE=1.65, p=.518, .95CI[-4.29, 2.16]. These results indicate that for each 1-point increase in perceptions of dominance, individuals showed a 2.84 Hz decline in pitch per utterance (for additional robustness checks see the OSM, Tables S8 and S9). Taken together, these analyses indicate that the effects of pitch modulation on emergent social rank occur primarily through galvanizing and sustaining perceptions of dominance—rooted in fear and intimidation—and not via altering perceptions of prestige.

Study 2: The Causal Impact of Pitch Modulation on Perceptions of Rank-Seeking Intentions

Like all naturalistic observational studies relying on a correlational approach, the conclusions drawn from Study 1 are limited by the possibility of third-factor variables influencing observed effects. In addition, although these results are suggestive of a causal effect of change in pitch on emergent rank outcomes, given that pitch changes were assessed in the initial moments of the interaction and rank was assessed at its conclusion it remains possible that rank asymmetries emerged and stabilized very early-on, such that our results instead capture the ways in which individuals modulate their pitch in response to appraisals of their current standing (Puts et al., 2006). To more directly test whether pitch alterations causally influence dominance perceptions and emergent social rank, in Study 2 we systematically manipulated the pitch trajectory of a vocal stimulus, and tested the causal impact of diverging pitch profiles on listeners' perceptions of vocalizers' rank-seeking motivations and dominance.

Method

Participants and Procedure. Two-hundred, seventy-four participants (60.58% women) participated in an online experiment. These individuals were recruited from two sources: an undergraduate subject pool at a large Canadian university (n = 181) who completed the experiment for course credit, and the online labor market pool Amazon Mechanical Turk (n = 93), who received monetary payment. Of note, prior studies deploying a similar perceptual task (i.e., using voice recordings) have found similar effects across online and offline (in the laboratory) presentation formats (Feinberg, DeBruine, Jones, & Perrett, 2008). Participants' age ranged from 15 to 61 (M = 23.62, SD = 7.64), and none reported any hearing impairments or difficulty. To ensure that all participants could perform the listening task, they were required to first listen to a test audio track and correctly answer a series of questions about this recording. Participants who failed to correctly answer any test question were excluded from participation.

Following a standard forced-choice playback paradigm used extensively in prior research on voice judgments (R. C. Anderson & Klofstad, 2012; Apicella & Feinberg, 2009; Tigue et al., 2012), participants listened to two vocal stimuli that differed in pitch trajectory (deepening or rising) but were otherwise identical. After listening to both recordings for as many times as they wanted, participants chose which of the pair of voices they considered more descriptive of a series of eight traits. Specifically, to assess perceptions of rank-seeking intentions, participants were asked to indicate which voice appeared more assertive, and which sought more power, leadership, and control. To assess perceptions of dominance, participants indicated which voice appeared more intimidating and threatening. Finally, to assess perceptions of prestige, participants indicated which voice appeared more admirable and respectable. Order of the recordings and questions presented were counterbalanced across participants (see OSM).

Experimental Stimuli. To create a "master" voice recording from which two stimuli with diverging pitch trajectories were subsequently generated, a male research assistant read, into a microphone, three scripted statements from a fictional work conversation (see OSM for details).

The master recording was finalized by standardizing the pitch of all three statements independently to 105 Hz. We then created two versions by manipulating the pitch of the first and last segment, without modifying the middle segment. In one version, the voice was altered to progressively deepen over the three statements. Specifically, the first segment was raised, and the last segment lowered, by 0.5 equivalent rectangular bandwidth (ERB), which is roughly equivalent to a perceived shift of ± 20 Hz; a change that is readily discernible and widely used in studies of pitch perceptions (Feinberg, Jones, Little, et al., 2005; Klofstad et al., 2012; Puts, Barndt, Welling, Dawood, & Burriss, 2011). In the other version, the voice was altered to progressively rise over the three statements. Specifically, the first segment was lowered, and the last segment raised, by 0.5 ERB. The three statements within each version were compiled into a single continuous recording.

The resultant pair of stimuli thus contained identical verbal content and virtually the same mean pitch across the entire recording, but differed in whether the pitch progressively deepened or rose across the three segments. All pitch manipulations were made independent of other acoustical properties (see OSM).

Analytical Approach. For all trait perceptions assessed, responses were coded ‘0’ for choice of the deepening pitch recording, and ‘1’ for the rising pitch recording. To derive an overall measure of perceived desire for social rank, we averaged the coded responses across the four relevant traits, which were highly inter-correlated (tetrachoric *r*s ranged from .69 to .78) and scaled together reliably (inter-item α=.81; M=.41, SD=.39). Responses on the two dominance items (tetrachoric r=.96) were similarly aggregated (inter-item α=.90; M=.44, SD=.47), as were the responses on the two prestige items (tetrachoric r=.70; inter-item α=.66; M=.53; SD=.43).

These three aggregate variables provide a summary measure of listeners’ tendency to perceive greater rank-seeking, dominance, and prestige in the recording with a deepening pitch profile, compared to the recording with a rising pitch profile. Lower values indicate a tendency to choose the deepening pitch recording as more descriptive of these traits, and higher values indicate a tendency to choose the rising pitch recording. Single sample Wilcoxon’s signed-rank tests were used to test the mean perception indices across the entire sample against 0.50, which would be the summary index expected if individuals randomly chose between the deepening and rising voices.

Results and Discussion. As predicted, a Wilcoxon signed-rank test indicated that listeners judged the deepening voice as expressing a greater desire for social rank than the rising voice (z=-3.55, p=.0004). Separate analyses using each of the four discrete items that comprise this measure revealed the same pattern; the deepening voice was more likely to be seen as assertive (z=-2.30, p=.022), and as seeking power (z=-3.26, p=.001), leadership (z=-2.90, p=.004), and control (z=-3.02, p=.003), relative to the rising voice. The size of these effects is comparable to previously documented effects of stable (i.e., rather than shifting) vocal pitch on leadership perceptions, using forced-choice playback designs (Klofstad, in press; Klofstad et al., 2012)

Turning to perceptions of dominance and prestige, we found that listeners judged the deepening voice as more dominant (z=-2.22, p=.027 than the rising voice). Also as predicted, perceptions of prestige were not influenced by pitch profiles (z=.98, p=.327; see Figure S2). These analyses were robust to the inclusion of controls to account for possible differences in participant gender, age, method of recruitment (student vs. Mechanical Turk), and stimuli order; ps=.035 and .795 (see OSM, Table S10), though the effect of manipulated pitch on perceived rank, despite maintaining a comparable effect size, became only marginally significant when these controls were applied (p=.115). See the OSM for details and additional analyses.

General Discussion

The present results support the prediction that humans signal rank intentions partly through systematic changes in vocal pitch, and these changes shape perceptions of relative formidability and the outcome of rank contests. By combining an externally valid observational study with an internally valid experiment manipulating perceived pitch alterations, the present research both examined the impact of dynamic changes in pitch as they actually occur in real-world rank contexts, and provided evidence for the causal impact of these alterations on rank outcomes. More specifically, Study 1 revealed that individuals interacting in face-to-face groups show spontaneous variation in the magnitude and direction of their pitch alterations, and these alterations predict their likelihood of winning or losing rank contests, such that individuals who raise their pitch over the initial moments of an interaction emerge as low-ranking, whereas those who deepen their pitch come to occupy higher ranks. Study 2 extended these findings by confirming that vocal changes, even in the absence of other accompanying behavioral signals, can causally influence perceptions of formidability and domination.

The present findings converge with and expand upon prior research on the use of pitch modulations to strategically signal self-assessed formidability. Puts and colleagues (2006) found that men who considered themselves physically stronger lowered their pitch when addressing a potential adversary, whereas those who considered themselves physically weaker raised their pitch. By broadening the scope of our investigation to the effects of this behavioral trait on the senders' emergent social rank and receivers' perceptions of these individuals, our results extend this prior work by demonstrating that strategic pitch changes—which are best interpreted as one of many evolved cues of dominance used by senders and receivers to readily establish relative formidability—systematically predict listeners' perceptions and actual rank outcomes.

Theoretical Implications

These findings are consistent with the notion that humans evolved to deploy dynamic vocal signals, along with a diverse and broader suite of cost-minimizing behavioral cues, that function to signal one's willingness to inflict costs on competitors, and consequently to enhance or dampen the signaler's apparent threat (Dovidio et al., 1988; Fournier et al., 2002; Marsh et al., 2014; Tiedens & Fragale, 2003). Moreover, these results are consistent with the expectation that our species is equipped with a specialized battery of perceptual biases designed to extract information from these vocal signals, to accurately estimate a rival's competitive motivation and relative formidability, and facilitate advantageous decisions about whether to defer or escalate in conflicts.

This evolutionary explanation for the observed effects contrasts with proximal explanations that focus on culturally acquired and reinforced gender associations and stereotypes. The cultural association hypothesis emphasizes that deference towards rivals who speak in lower voices and display other masculine cues derives strictly from cultural beliefs, institutions, and practices that confer status to men or individuals who exhibit masculine traits, without any functional basis. One implication of this view is that in other societies vocal pitch may not be associated with dominance at all, or deeper pitch voices might signal submission or shame. Two aspects of our findings, however, challenge this expectation. First, our results indicate that lowering pitch enhances dominance but not prestige—whereas the latter would be expected, as well, from a cultural association account, given cultural beliefs linking men to competence and success (Eagly & Mladinic, 1994).

Second, we found that both men and women spontaneously deploy pitch modulation signals to effectively exaggerate their formidability, suggesting that this is not only a signal for men. Furthermore, recent evidence suggests that the dominance avenue to social rank operates effectively in both genders, and influences group decision-making, deference, and attention (Cheng et al., 2013, 2010b; Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010). Although men and women may ultimately differ in their use of numerous specific tactics to effectively advertise and exaggerate their threat and formidability in efforts to prevail in dominance disputes—with males' preferences tailored towards direct and aggressive physical combat, in contrast to females' more indirect and low-cost aggressive tactics (e.g., shunning, gossiping, ostracizing; Archer, 2004; Campbell, 1999)—our findings suggest that within the vocal domain, pitch modulation signals operate similarly across both genders to augment (or diminish) dominance rank. These results are consistent with prior experimental work showing that lower pitched female (and male) voices are deemed more dominant by listeners of both genders (Jones et al., 2010), and that female listeners tend to be more sensitive to this dominance cue in female voices, and reveal a stronger perceptual bias than male listeners (Borkowska & Pawlowski, 2011). Similarly, in electing political leaders, men and women not only prefer female candidates with lower voices—even when the contested leadership role is feminine in nature (e.g., President of the Parent Teachers Organization; R. C. Anderson & Klofstad, 2012)—but this preference is stronger when evaluating female candidates than male candidates (Klofstad, in press). In light of these prior results, it is not surprising that we found the same qualitative relationship between a deepening pitch profile and greater dominance and social rank among both men and women.

Moreover, evidence from two other lines of research, which converge with the evolutionary-based explanation, also challenges the cultural association explanation. First, a strict cultural learning explanation cannot readily account for the well-documented pitch- and pitch modulation-regulated patterns of dominance and submission observed across a wide range of non-human animals (Bee et al., 1999; Davies & Halliday, 1978; Morton, 1977; Morton & Page, 1992; Owings & Morton, 1998; Ryan & Brenowitz, 1985; Vannoni & McElligott, 2008; Wagner Jr., 1992), and in evolutionarily relevant small-scale societies (Apicella, Feinberg, & Marlowe, 2007; Hodges-Simeon et al., 2014; Puts et al., 2012). The Hadza hunter-gatherers in Tanzania are particularly relevant here, as they are known for their relatively gender egalitarian social norms and matrilocal biases. Despite these features of Hadza society, a deeper vocal pitch is associated with higher status among these individuals (Apicella & Feinberg, 2009; Apicella et al., 2007). This commonality across species and human societies is consistent with the notion that submission to individuals with deeper pitched voices arises primarily from an evolved psychology.

Second, it is not clear how the cultural association explanation might account for emerging evidence of a biological underpinning of vocal pitch, and its effects on mate choice.

For example, studies suggest that higher levels of testosterone are linked to low pitch in men (Bruckert et al., 2006; Dabbs Jr. & Mallinger, 1999; Harries et al., 1998; Puts et al., 2012), and women show a stronger preference for lower pitched men during the fertile period of their ovulatory cycle (Feinberg et al., 2006; Puts, 2005), non-lactating phases (Apicella & Feinberg, 2009), and in short-term mating contexts (Puts, 2005). These findings are consistent with—and predicted *a priori* by—the evolutionary approach, which proposes that low vocal pitch and other masculine displays (e.g., facial masculinity, body symmetry) function as androgen dependent traits that signal threat potential, resource acquisition capacity, and mate quality (Andersson, 1994; Apicella et al., 2007; Feinberg, 2008; Gangestad, Simpson, Cousins, Garver-Apgar, & Christensen, 2004; Penton-Voak & Perrett, 2000; Puts et al., 2012). In contrast, it is not clear why and how cultural institutions and systems alone would give rise to women's capacity to discern among these features and to use them strategically to guide their mate choice under different ovulatory and lactation statuses. Nonetheless, firm conclusions about whether these dominance signals represent evolved aspects of human psychology, and the role of cultural traditions and expectations in shaping or increasing the influence of pitch on social perceptions, await future work, ideally including studies that examine the generalizability and variability of these signals across a diverse range of environments and populations.

The present results also bear implications for signaling theory, an extensive theoretical program aimed at understanding how selection has shaped behavioral, physiological, and morphological characteristics designed to facilitate communication and informational exchange between organisms (Bliege Bird & Smith, 2005; Maynard Smith & Harper, 1995; Otte, 1974; Rendall, Owren, & Ryan, 2009). Central to this enterprise is the concept of signal honesty, which considers the degree to which a signal accurately (though not necessarily perfectly) encodes information about the signaler's underlying qualities, such as size, condition, or intention (Maynard Smith & Harper, 2003; Searcy & Nowicki, 2005; Zahavi, 1975; Zahavi & Zahavi, 1997). Does vocal pitch provide an honest signal of an individual's formidability? The evidence, across different formidability-relevant characteristics and traits, appears mixed. With regard to body size and strength—the primary physical determinants of fighting potential across species, including humans (Archer, 1988)—the answer may be no; studies have generally shown weak to null associations between pitch and physical size and upper-body strength, within sexes (Collins & Missing, 2003; Rendall, Vokey, & Nemeth, 2007; Sell et al., 2010; but see Evans et al., 2006; Puts et al., 2012). On the other hand, there is considerable evidence linking vocal pitch to testosterone, which is associated with physical aggressiveness and hostility (Archer, 1991, 2006; Aromäki, Lindman, & Eriksson, 1999; Book, Starzyk, & Quinsey, 2001), consistent with the notion that deep voices serve as an honest advertisement of threat potential (Bruckert et al., 2006; Dabbs Jr. & Mallinger, 1999; Evans et al., 2008; Puts et al., 2012).

With regard to threat intentions, the present results, combined with other evidence linking deepening voices and self-appraised or imminent dominance (e.g., attacking) in humans (Gregory Jr. et al., 2001; Gregory Jr. & Webster, 1996; Puts et al., 2006), tentatively suggests that a lowering pitch may honestly signal a motivation and readiness to vie for dominance in rank competitions. Beyond advertising competitive intentions, however, the extent to which this signal correlates with the producer's strength or other formidability enhancing qualities to also function as an honest signal of fighting potential, is an interesting question for future inquiry. Future work should also further explore the effects of pitch changes on perceivers' behaviors, such as patterns of deference, attention, and escalation.

Limitations and Future Directions

These findings open up several avenues for potentially fruitful future work. One area concerns the vocal dynamics that underpin rank contests in women. For example, given that we examined the causal effects of pitch change in a male voice only, subsequent work should directly address whether listeners also readily extract dominance information from female speakers' pitch modulations. The results of Study 1, which indicate a highly consistent pattern of effects across female and male groups, ameliorate this concern to some extent, but future research is needed to examine the effects of pitch alterations in mixed-gender contests. In fact, one prior study hints at the possibility that lowering versus rising pitch signals may have differential effects on rank attainment in competitions against rivals of the opposite sex (Klofstad, in press), but more work is needed in this area.

Another area ripe for future research involves the role of situational changes in vocal pitch in the related domain of mating preferences. A growing number of studies suggest that women have a generalized preference for deeper male voices (Collins, 2000; Feinberg et al., 2006; Feinberg, Jones, Little, et al., 2005; Hodges-Simeon, Gaulin, & Puts, 2010), and a study examining the aforementioned matrilocally inclined hunter-gatherers found that men in this society with lower-pitched voices have greater reproductive success (Apicella et al., 2007). In contrast, men have a generalized preference for higher-pitched female voices (Feinberg, Jones, DeBruine, et al., 2005; Feinberg et al., 2008), which (like other expressions of femininity) signal reproductive health and fertility, hormonal profile, and age (Bryant & Haselton, 2009; Collins & Missing, 2003; Thornhill & Gangestad, 1999). While these existing studies link stable between-person variation in vocal pitch to differential mate value and attractiveness, more research is needed to examine how within-person vocal changes operate in mating contexts, and potentially covary with hormonal variation and ovulation cycles (Amir & Biron-Shental, 2004; Bryant & Haselton, 2009).

An additional important direction for future work entails examining whether and how pitch changes operate in zero-sum competitions, and whether effects differ from those observed in the collaborative team environment examined here. One possibility is that, in zero-sum, competitive contests, a deepening voice may be used to cue one's competitive intentions and likelihood of success. Research on hormone changes suggests that winners of skill-based contests devoid of agonistic conflicts (e.g., chess matches, cognitive competitions involving puzzles) experience post-victory momentary surges in testosterone (Mazur, Booth, & Dabbs, 1992; Mehta & Josephs, 2006; Zilioli & Watson, 2012), which may in turn produce a corresponding drop in vocal pitch (Harries et al., 1998). These findings raise the possibility that, in contrast to collaborative team environments in which knowledgeable and capable individuals should suppress deepening vocal signals so as to avoid appearing dominant and risk losing prestige in the eyes of one's group members, in zero-sum exchanges the relatively more successful (prestigious) competitor may emit pitch changes that cue his/her entry into a competitive state.

Finally, several methodological limitations of the present research should be addressed in future work. First is the absence of external incentives (e.g., experimental rewards or punishments) provided for rank attainment (to be clear, participants in our study were given a team incentive for correct responses on the task, but there was no incentive for group members to compete against each other for any kind of rank or status). Even in the ecologically valid situation of completing a group task in Study 1, participants might not have been strongly motivated to pursue high rank; the rewards of attaining high rank within contrived laboratory groups may not be entirely clear, and could be considered largely psychological. However, these psychological benefits should not be underestimated. Converging lines of behavioral, cross-species, and neuro-scientific evidence suggest that attaining high rank is intrinsically rewarding, motivating, and universal (C. Anderson, Hildreth, & Howland, 2015; Martinez et al., 2010; Tamashiro, Nguyen, & Sakai, 2005). Nonetheless, one important future research direction is to examine vocal signal dynamics when individuals are directly incentivized to acquire high rank, in a wide range of contexts and groups.

A second methodological limitation concerns our measure of pitch, which likely provides a rough and limited proxy for the full range and extent of the actual vocal changes produced, as pitch was derived from assessing only three vocal samples rather than over a longer time span. We expect that the actual pitch change signals we are ultimately seeking to understand in dominance contests are more salient and exaggerated than those documented here. Future work should examine vocalizations generated over the full duration of interactions to establish the precise magnitude and strength of pitch change signals displayed in rank contests, and employ psychoacoustic studies to assess the perceptual relevance of these (possibly larger) naturally occurring pitch changes.

Despite these limitations and important future research directions, the present findings provide strong support for the suggestion that: (a) pitch alteration signals are produced spontaneously and strategically in real face-to-face rank contests; (b) receivers both detect pitch changes and use the reliable information provided by these signals to accurately gauge rank intentions; and (c) these vocal signals influence senders’ perceived formidability, and predict who prevails and who submits in a contest. Combined with prior work, our findings are consistent with the idea that humans rely on systematic vocal changes, alongside a broad suite of other evolved behavioral signals, to track, signal and coordinate hierarchical relationships.

References

Adler, N. E., Boyce, T., Chesney, M. A., Cohen, S., Folkman, S., Kahn, R. L., & Leonard, S. (1994). Socioeconomic status and health: The challenge of the gradient. *American Psychologist*, *49*(1), 15–24. http://doi.org/10.1037/0003-066X.49.1.15

Amir, O., & Biron-Shental, T. (2004). The impact of hormonal fluctuations on female vocal folds. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *12*(3), 180–184.

Anderson, C., Hildreth, J. A. D., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological Bulletin*, *141*(3), 574–601. http://doi.org/10.1037/a0038781

Anderson, R. C., & Klofstad, C. A. (2012). Preference for Leaders with Masculine Voices Holds in the Case of Feminine Leadership Roles. *PLoS ONE*, *7*(12), e51216. http://doi.org/10.1371/journal.pone.0051216

Andersson, M. B. (1994). *Sexual selection*. Princeton, NJ: Princeton University Press.

Anicich, E. M., Swaab, R. I., & Galinsky, A. D. (2015). Hierarchical cultural values predict success and mortality in high-stakes teams. *Proceedings of the National Academy of Sciences*, *112*(5), 1338–1343. http://doi.org/10.1073/pnas.1408800112

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1659), 1077–1082. http://doi.org/10.1098/rspb.2008.1542

Apicella, C. L., Feinberg, D. R., & Marlowe, F. W. (2007). Voice pitch predicts reproductive success in male hunter-gatherers. *Biology Letters*, *3*(6), 682–684. http://doi.org/10.1098/rsbl.2007.0410

Archer, J. (1988). *The behavioural biology of aggression*. Cambridge, UK: Cambridge University Press.

Archer, J. (1991). The influence of testosterone on human aggression. *British Journal of Psychology*, *82*(1), 1–28. http://doi.org/10.1111/j.2044-8295.1991.tb02379.x

Archer, J. (2004). Sex differences in aggression in real-world settings: A meta-analytic review. *Review of General Psychology*, *8*(4), 291–322. http://doi.org/10.1037/1089-2680.8.4.291

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews*, *30*(3), 319–345. http://doi.org/10.1016/j.neubiorev.2004.12.007

Aromäki, A. S., Lindman, R. E., & Eriksson, C. J. P. (1999). Testosterone, aggressiveness, and antisocial personality. *Aggressive Behavior*, *25*(2), 113–123. http://doi.org/10.1002/(SICI)1098-2337(1999)25:2<113::AID-AB4>3.0.CO;2-4

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, *70*(3), 614.

Bass, B. M. (1954). The leaderless group discussion. *Psychological Bulletin*, *51*(5), 465–492. http://doi.org/10.1037/h0056881

Bee, M. A., Perrill, S. A., & Owen, P. C. (1999). Size assessment in simulated territorial encounters between male green frogs (Rana clamitans). *Behavioral Ecology and Sociobiology*, *45*(3-4), 177–184. http://doi.org/10.1007/s002650050551

Bee, M. A., Perrill, S. A., & Owen, P. C. (2000). Male green frogs lower the pitch of acoustic signals in defense of territories: A possible dishonest signal of size? *Behavioral Ecology*, *11*(2), 169–177. http://doi.org/10.1093/beheco/11.2.169

Bendersky, C., & Hays, N. A. (2012). Status conflict in groups. *Organization Science, 23*(2), 323–340. http://doi.org/10.1287/orsc.1110.0734

Bliege Bird, R., & Smith, E. A. (2005). Signaling theory, strategic interaction, and symbolic capital. *Current Anthropology*, *46*(2), 221–248. http://doi.org/10.1086/427115

Book, A. S., Starzyk, K. B., & Quinsey, V. L. (2001). The relationship between testosterone and aggression: A meta-analysis. *Aggression and Violent Behavior*, *6*(6), 579–599. http://doi.org/10.1016/S1359-1789(00)00032-X

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59.

Bottger, P. C. (1984). Expertise and air time as bases of actual and perceived influence in problem-solving groups. *Journal of Applied Psychology*, *69*(2), 214–221. http://doi.org/10.1037/0021-9010.69.2.214

Bro-Jørgensen, J. (2010). Dynamics of multiple signalling systems: Animal communication in a world in flux. *Trends in Ecology & Evolution*, *25*(5), 292–300. http://doi.org/10.1016/j.tree.2009.11.003

Brown, D. E. (1991). *Human universals*. New York, New York: McGraw-Hill.

Bruckert, L., Liénard, J.-S., Lacroix, A., Kreutzer, M., & Leboucher, G. (2006). Women use voice parameters to assess men's characteristics. *Proceedings of the Royal Society B: Biological Sciences*, *273*(1582), 83–89. http://doi.org/10.1098/rspb.2005.3265

Bryant, G. A., & Haselton, M. G. (2009). Vocal cues of ovulation in human females. *Biology Letters*, *5*(1), 12–15. http://doi.org/10.1098/rsbl.2008.0507

Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin, 101*(1), 147–158. http://doi.org/10.1037/0033-2909.101.1.147

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*(01), 1–14. http://doi.org/10.1017/S0140525X00023992

Campbell, A. (1999). Staying alive: Evolution, culture, and women's intrasexual aggression. *Behavioral and Brain Sciences*, *22*(02), 203–214. http://doi.org/null

Carré, J. M., & McCormick, C. M. (2008). In your face: facial metrics predict aggressive behaviour in the laboratory and in varsity and professional hockey players. *Proceedings of the Royal Society B: Biological Sciences*, *275*(1651), 2651–2656. http://doi.org/10.1098/rspb.2008.0873

Chance, M. R. A. (1967). Attention Structure as the Basis of Primate Rank Orders. *Man*, *2*(4), 503–518. http://doi.org/10.2307/2799336

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. http://doi.org/10.1037/a0030398

Cheng, J. T., Tracy, J. L., & Henrich, J. (2010a). Pride, personality, and the evolutionary foundations of human social status. *Evolution and Human Behavior*, *31*(5), 334–347. http://doi.org/10.1016/j.evolhumbehav.2010.02.004

Cheng, J. T., Tracy, J. L., & Henrich, J. (2010b). Pride, personality, and the evolutionary foundations of human social status. *Evolution and Human Behavior*, *31*(5), 334–347. http://doi.org/10.1016/j.evolhumbehav.2010.02.004

Christian, J. J. (1970). Social subordination, population density, and mammalian evolution. *Science*, *168*(3927), 84–90. http://doi.org/10.1126/science.168.3927.84

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, *60*(6), 773–780.

Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour*, *65*(5), 997–1004. http://doi.org/10.1006/anbe.2003.2123

Curhan, J. R., & Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. *Journal of Applied Psychology*, *92*(3), 802–811. http://doi.org/10.1037/0021-9010.92.3.802

Dabbs Jr., J. M., & Mallinger, A. (1999). High testosterone levels predict low voice pitch among men. *Personality and Individual Differences*, *27*(4), 801–804.

Davies, N. B., & Halliday, T. R. (1978). Deep croaks and fighting assessment in toads Bufo bufo. *Nature*, *274*(5672), 683–685. http://doi.org/10.1038/274683a0

de Waal, F. B. M. (1982). *Chimpanzee politics*. New York, NY: Harper & Row.

de Waal, F. B. M., & Luttrell, L. M. (1985). The formal hierarchy of rhesus macaques: An investigation of the bared-teeth display. *American Journal of Primatology*, *9*(2), 73–85. http://doi.org/10.1002/ajp.1350090202

Dovidio, J. F., Ellyson, S. L., Keating, C. F., Heltman, K., & Brown, C. E. (1988). The relationship of social power to visual displays of dominance between men and women.

Journal of Personality and Social Psychology, *54*(2), 233–242. http://doi.org/10.1037/0022-3514.54.2.233

Eagly, A. H., & Mladinic, A. (1994). Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *European Review of Social Psychology*, *5*(1), 1–35. http://doi.org/10.1080/14792779543000002

Eisenegger, C., Haushofer, J., & Fehr, E. (2011). The role of testosterone in social interaction. *Trends in Cognitive Sciences*, *15*(6), 263–271. http://doi.org/10.1016/j.tics.2011.04.008

Evans, S., Neave, N., & Wakelin, D. (2006). Relationships between vocal characteristics and body size and shape in human males: An evolutionary explanation for a deep male voice. *Biological Psychology*, *72*(2), 160–163.

Evans, S., Neave, N., Wakelin, D., & Hamilton, C. (2008). The relationship between testosterone and vocal frequencies in human males. *Physiology & Behavior*, *93*(4), 783–788.

Feinberg, D. R. (2008). Are human faces and voices ornaments signaling common underlying cues to mate value? *Evolutionary Anthropology: Issues, News, and Reviews*, *17*(2), 112–118. http://doi.org/10.1002/evan.20166

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception*, *37*(4), 615–623. http://doi.org/10.1068/p5514

Feinberg, D. R., Jones, B. C., DeBruine, L. M., Moore, F. R., Law Smith, M. J., Cornwell, R. E., … Perrett, D. I. (2005). The voice and face of woman: One ornament that signals quality? *Evolution and Human Behavior*, *26*(5), 398–408. http://doi.org/10.1016/j.evolhumbehav.2005.04.001

Feinberg, D. R., Jones, B. C., Law Smith, M. J., Moore, F. R., DeBruine, L. M., Cornwell, R. E., … Perrett, D. I. (2006). Menstrual cycle, trait estrogen level, and masculinity preferences in the human voice. *Hormones and Behavior*, *49*(2), 215–222. http://doi.org/10.1016/j.yhbeh.2005.07.004

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, *69*(3), 561–568. http://doi.org/10.1016/j.anbehav.2004.06.012

Fessler, D. M. T., Holbrook, C., & Snyder, J. K. (2012). Weapons make the man (larger): Formidability is represented as size and strength in humans. *PloS One*, *7*(4), e32751.

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, *117*(3), 319–331. http://doi.org/10.1016/j.cognition.2010.09.003

Fournier, M. A. (2009). Adolescent Hierarchy Formation and the Social Competition Theory of Depression. *Journal of Social and Clinical Psychology*, *28*(9), 1144–1172. http://doi.org/10.1521/jscp.2009.28.9.1144

Fournier, M. A., Moskowitz, D. S., & Zuroff, D. C. (2002). Social rank strategies in hierarchical relationships. *Journal of Personality and Social Psychology*, *83*(2), 425–433. http://doi.org/10.1037/0022-3514.83.2.425

Gangestad, S. W., Simpson, J. A., Cousins, A. J., Garver-Apgar, C. E., & Christensen, P. N. (2004). Women's preferences for male behavioral displays change across the menstrual cycle. *Psychological Science*, *15*(3), 203–207. http://doi.org/10.1111/j.0956-7976.2004.01503010.x

Gerhardt, H. C. (1994). The Evolution of Vocalization in Frogs and Toads. *Annual Review of Ecology and Systematics*, *25*, 293–324.

Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus monkey (Macaca mulatta) screams: Representational signalling in the recruitment of agonistic aid. *Animal Behaviour*, *32*(1), 182–193. http://doi.org/10.1016/S0003-3472(84)80336-X

Gregory Jr., S. W., Dagan, K., & Webster, S. (1997). Evaluating the relation of vocal accommodation in conversation partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior*, *21*(1), 23–43. http://doi.org/10.1023/A:1024995717773

Gregory Jr., S. W., Green, B. E., Carrothers, R. M., Dagan, K., & Webster, S. (2001). Verifying the primacy of voice fundamental frequency in social status accommodation. *Language & Communication*, *21*(1), 37–60. http://doi.org/10.1016/S0271-5309(00)00011-2

Gregory Jr., S. W., & Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, *70*(6), 1231–1240. http://doi.org/10.1037/0022-3514.70.6.1231

Halevy, N., Chou, E. Y., Galinsky, A. D., & Murnighan, J. K. (2012). When Hierarchy Wins Evidence From the National Basketball Association. *Social Psychological and Personality Science*, *3*(4), 398–406. http://doi.org/10.1177/1948550611424225

Harries, M., Hawkins, S., Hacking, J., & Hughes, I. (1998). Changes in the male voice at puberty: Vocal fold length and its relationship to the fundamental frequency of the voice. *The Journal of Laryngology & Otology*, *112*(05), 451–454. http://doi.org/10.1017/S0022215100140757

Hawley, P. H. (2002). Social dominance and prosocial and coercive strategies of resource control in preschoolers. *International Journal of Behavioral Development*, *26*(2), 167–176. http://doi.org/10.1080/01650250042000726

Henrich, J., Chudek, M., & Boyd, R. (2015). The Big Man Mechanism: how prestige fosters cooperation and creates prosocial leaders. *Phil. Trans. R. Soc. B*, *370*(1683), 20150013. http://doi.org/10.1098/rstb.2015.0013

Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*(3), 165–196. http://doi.org/10.1016/S1090-5138(00)00071-4

Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different vocal parameters predict perceptions of dominance and attractiveness. *Human Nature*, *21*(4), 406–427. http://doi.org/10.1007/s12110-010-9101-5

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. http://doi.org/10.1093/beheco/aru081

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *1*(79), 57–62. http://doi.org/10.1016/j.anbehav.2009.10.003

Judge, T. A., & Cable, D. M. (2004). The Effect of Physical Height on Workplace Success and Income: Preliminary Test of a Theoretical Model. *Journal of Applied Psychology*, *89*(3), 428–441. http://doi.org/10.1037/0021-9010.89.3.428

Kalma, A. (1991). Hierarchisation and dominance assessment at first glance. *European Journal of Social Psychology*, *21*(2), 165–181. http://doi.org/10.1002/ejsp.2420210206

Keltner, D., Gruenfeld, D. H., & Anderson, C. (2003). Power, approach, and inhibition. *Psychological Review*, *110*(2), 265–284. http://doi.org/10.1037/0033-295X.110.2.265

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford Press.

Klofstad, C. A. (in press). Candidate voice pitch influences election outcomes. *Political Psychology*. http://doi.org/10.1111/pops.12280

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. http://doi.org/10.1098/rspb.2012.0311

Ko, S. J., Sadler, M. S., & Galinsky, A. D. (2015). The Sound of Power Conveying and Detecting Hierarchical Rank Through Voice. *Psychological Science*, *26*(1), 3–14. http://doi.org/10.1177/0956797614553009

Ladefoged, P. (1996). *Elements of acoustic phonetics* (2nd ed.). Chicago, IL: University of Chicago Press.

Lee, R. B. (1979). *The! Kung San: Men, women, and work in a foraging society*. Cambridge, NY: Cambridge University Press.

Lewis, H. S. (1974). *Leaders and followers: Some anthropological perspectives*. Reading, MA: Addison-Wesley.

Lorenz, K. (1966). *On aggression*. New York, New York: Harcourt, Brace and World.

Marsh, A. A., Cardinale, E. M., Chentsova-Dutton, Y. E., Grossman, M. R., & Krumpos, K. A. (2014). Power Plays Expressive Mimicry of Valid Agonistic Cues. *Social Psychological and Personality Science*, *5*(6), 684–690. http://doi.org/10.1177/1948550613519684

Martinez, D., Orlowska, D., Narendran, R., Slifstein, M., Liu, F., Kumar, D., … Kleber, H. D. (2010). Dopamine type 2/3 receptor availability in the striatum and social status in human volunteers. *Biological Psychiatry*, *67*(3), 275–278. http://doi.org/10.1016/j.biopsych.2009.07.037

Maynard Smith, J., & Harper, D. G. C. (1995). Animal signals: Models and terminology. *Journal of Theoretical Biology*, *177*(3), 305–311. http://doi.org/10.1006/jtbi.1995.0248

Maynard Smith, J., & Harper, D. G. C. (2003). *Animal signals*. Oxford, UK: Oxford University Press.

Maynard Smith, J., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal Behaviour*, *24*(1), 159–175. http://doi.org/10.1016/S0003-3472(76)80110-8

Maynard Smith, J., & Price, G. R. (1973). The logic of animal conflict. *Nature*, *246*(5427), 15–18. http://doi.org/10.1038/246015a0

Mazur, A. (1985). A biosocial model of status in face-to-face primate groups. *Social Forces*, *64*(2), 377–402.

Mazur, A., Booth, A., & Dabbs, J. M. (1992). Testosterone and chess competition. *Social Psychology Quarterly*, *55*(1), 70–77. http://doi.org/10.2307/2786687

Mehta, P. H., & Josephs, R. A. (2006). Testosterone change after losing predicts the decision to compete again. *Hormones and Behavior*, *50*(5), 684–692. http://doi.org/10.1016/j.yhbeh.2006.07.001

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. *The American Naturalist*, *111*(981), 855–869.

Morton, E. S., & Page, J. (1992). *Animal talk: Science and the voices of nature*. New York, NY: Random House.

Otte, D. (1974). Effects and functions in the evolution of signaling systems. *Annual Review of Ecology and Systematics*, *5*, 385–417.

Owings, D. H., & Morton, E. S. (1998). *Animal vocal communication: A new approach*. Cambridge, UK: Cambridge University Press.

Parker, G. A. (1974). Assessment strategy and the evolution of fighting behaviour. *Journal of Theoretical Biology*, *47*(1), 223–243. http://doi.org/10.1016/0022-5193(74)90111-8

Penton-Voak, I. S., & Perrett, D. I. (2000). Female preference for male faces changes cyclically: Further evidence. *Evolution and Human Behavior*, *21*(1), 39–48. http://doi.org/10.1016/S1090-5138(99)00033-1

Puts, D. A. (2005). Mating context and menstrual phase affect women's preferences for male voice pitch. *Evolution and Human Behavior*, *26*(5), 388–397. http://doi.org/10.1016/j.evolhumbehav.2005.03.001

Puts, D. A., Apicella, C. L., & Cárdenas, R. A. (2012). Masculine voices signal men's threat potential in forager and industrial societies. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1728), 601–609. http://doi.org/10.1098/rspb.2011.0829

Puts, D. A., Barndt, J. L., Welling, L. L., Dawood, K., & Burriss, R. P. (2011). Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences*, *50*(1), 111–115.

Puts, D. A., Gaulin, S. J. C., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, *27*(4), 283–296. http://doi.org/10.1016/j.evolhumbehav.2005.11.003

Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, *28*(5), 340–344.

Rendall, D., Owren, M. J., & Ryan, M. J. (2009). What do animal signals mean? *Animal Behaviour*, *78*(2), 233–240. http://doi.org/10.1016/j.anbehav.2009.06.007

Rendall, D., Vokey, J. R., & Nemeth, C. (2007). Lifting the curtain on the Wizard of Oz: Biased voice-based impressions of speaker size. *Journal of Experimental Psychology: Human Perception and Performance; Journal of Experimental Psychology: Human Perception and Performance*, *33*(5), 1208–1219.

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability the difference score in the measurement of change. *Journal of Educational Measurement*, *20*(4), 335–343. http://doi.org/10.1111/j.1745-3984.1983.tb00211.x

Ronay, R., Greenaway, K., Anicich, E. M., & Galinsky, A. D. (2012). The path to glory Is paved with hierarchy: When hierarchical differentiation increases group effectiveness. *Psychological Science*, *23*(6), 669–677.

Ryan, M. J., & Brenowitz, E. A. (1985). The role of body size, phylogeny, and ambient noise in the evolution of bird song. *The American Naturalist*, *126*(1), 87–100.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, *99*(2), 143–165. http://doi.org/10.1037/0033-2909.99.2.143

Schjelderup-Ebbe, T. (1922). Beiträge zur sozialpsychologie des haushuhns [Contributions to the social psychology of the domestic chicken]. *Zeitschrift Für Psychologie*, *88*, 225–252.

Searcy, W. A., & Nowicki, S. (2005). *The evolution of communication: Reliability and deception in animal signaling systems*. Princeton, NJ: Princeton University Press.

Sell, A., Bryant, G. A., Cosmides, L., Tooby, J., Sznycer, D., Von Rueden, C., … Gurven, M. (2010). Adaptations in humans for assessing physical strength from the voice.

Proceedings of the Royal Society B: Biological Sciences, *277*, 3509–3518. http://doi.org/10.1098/rspb.2010.0769

Sell, A., Cosmides, L., Tooby, J., Sznycer, D., von Rueden, C., & Gurven, M. (2009). Human adaptations for the visual assessment of strength and fighting ability from the body and face. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1656), 575–584. http://doi.org/10.1098/rspb.2008.1177

Shariff, A. F., & Tracy, J. L. (2009). Knowing who’s boss: Implicit perceptions of status from the nonverbal expression of pride. *Emotion*, *9*(5), 631–639. http://doi.org/10.1037/a0017089

Shariff, A. F., Tracy, J. L., & Markusoff, J. L. (2012). (Implicitly) Judging a Book by Its Cover The Power of Pride and Shame Expressions in Shaping Judgments of Social Status. *Personality and Social Psychology Bulletin*, *38*(9), 1178–1193. http://doi.org/10.1177/0146167212446834

Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. New York, NY: Cambridge University Press.

Sinnott, J. M., Owren, M. J., & Petersen, M. R. (1987). Auditory frequency discrimination in primates: Species differences (Cercopithecus, Macaca, Homo). *Journal of Comparative Psychology*, *101*(2), 126–131. http://doi.org/10.1037/0735-7036.101.2.126

Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences*, *104*(43), 17228–17233. http://doi.org/10.1073/pnas.0706741104

Smith, D. R. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T. (2005). The processing and perception of size information in speech sounds. *The Journal of the Acoustical Society of America*, *117*(1), 305–318.

Tamashiro, K. L. K., Nguyen, M. M. N., & Sakai, R. R. (2005). Social stress: From rodents to primates. *Frontiers in Neuroendocrinology*, *26*(1), 27–40. http://doi.org/10.1016/j.yfrne.2005.03.001

Thomsen, L., Frankenhuis, W. E., Ingold-Smith, M., & Carey, S. (2011). Big and mighty: Preverbal infants mentally represent social dominance. *Science*, *331*(6016), 477–480. http://doi.org/10.1126/science.1199198

Thornhill, R., & Gangestad, S. W. (1999). Facial attractiveness. *Trends in Cognitive Sciences*, *3*(12), 452–460. http://doi.org/10.1016/S1364-6613(99)01403-5

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology*, *84*(3), 558–568. http://doi.org/10.1037/0022-3514.84.3.558

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. http://doi.org/10.1016/j.evolhumbehav.2011.09.004

Tinbergen, N. (1959). Behaviour, systematics, and natural selection. *Ibis*, *101*(3-4), 318–330. http://doi.org/10.1111/j.1474-919X.1959.tb02388.x

Tracy, J. L., Shariff, A. F., Zhao, W., & Henrich, J. (2013). Cross-cultural evidence that the nonverbal expression of pride is an automatic status signal. *Journal of Experimental Psychology: General*, *142*(1), 163–180. http://doi.org/10.1037/a0028412

Vannoni, E., & McElligott, A. G. (2008). Low frequency groans indicate larger and more dominant fallow deer (Dama dama) males. *PLoS ONE*, *3*(9), e3113. http://doi.org/10.1371/journal.pone.0003113

von Rueden, C., Gurven, M., & Kaplan, H. (2011). Why do men seek status? Fitness payoffs to dominance and prestige. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1715), 2223.

Wagner Jr., W. E. (1992). Deceptive or honest signalling of fighting ability? A test of alternative hypotheses for the function of changes in call dominant frequency by male cricket frogs. *Animal Behaviour*, *44*, 449–462. http://doi.org/10.1016/0003-3472(92)90055-E

Wolff, S. E., & Puts, D. A. (2010). Vocal masculinity is a robust dominance signal in men. *Behavioral Ecology and Sociobiology*, *64*(10), 1673–1683.

Zahavi, A. (1975). Mate selection—A selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214. http://doi.org/10.1016/0022-5193(75)90111-3

Zahavi, A., & Zahavi, A. (1997). *The handicap principle: A missing piece of Darwins puzzle*. Oxford, UK: Oxford University Press.

Zilioli, S., & Watson, N. V. (2012). The hidden dimensions of the competition effect: Basal cortisol and basal testosterone jointly predict changes in salivary testosterone after social victory in men. *Psychoneuroendocrinology*, *37*(11), 1855–1865. http://doi.org/10.1016/j.psyneuen.2012.03.022

Table Caption

Table 1. Hierarchical linear model predicting vocal pitch from the main and interactive effects of time and rank, controlling for gender and its interaction with time. Coefficients are followed by standard errors in parentheses. Results indicate that social standing significantly predicts vocal pitch trajectories, in the form of a time × social rank interaction; a 1-point increase in social rank is associated with a 4.45 Hz drop in pitch per utterance.

Figure Caption

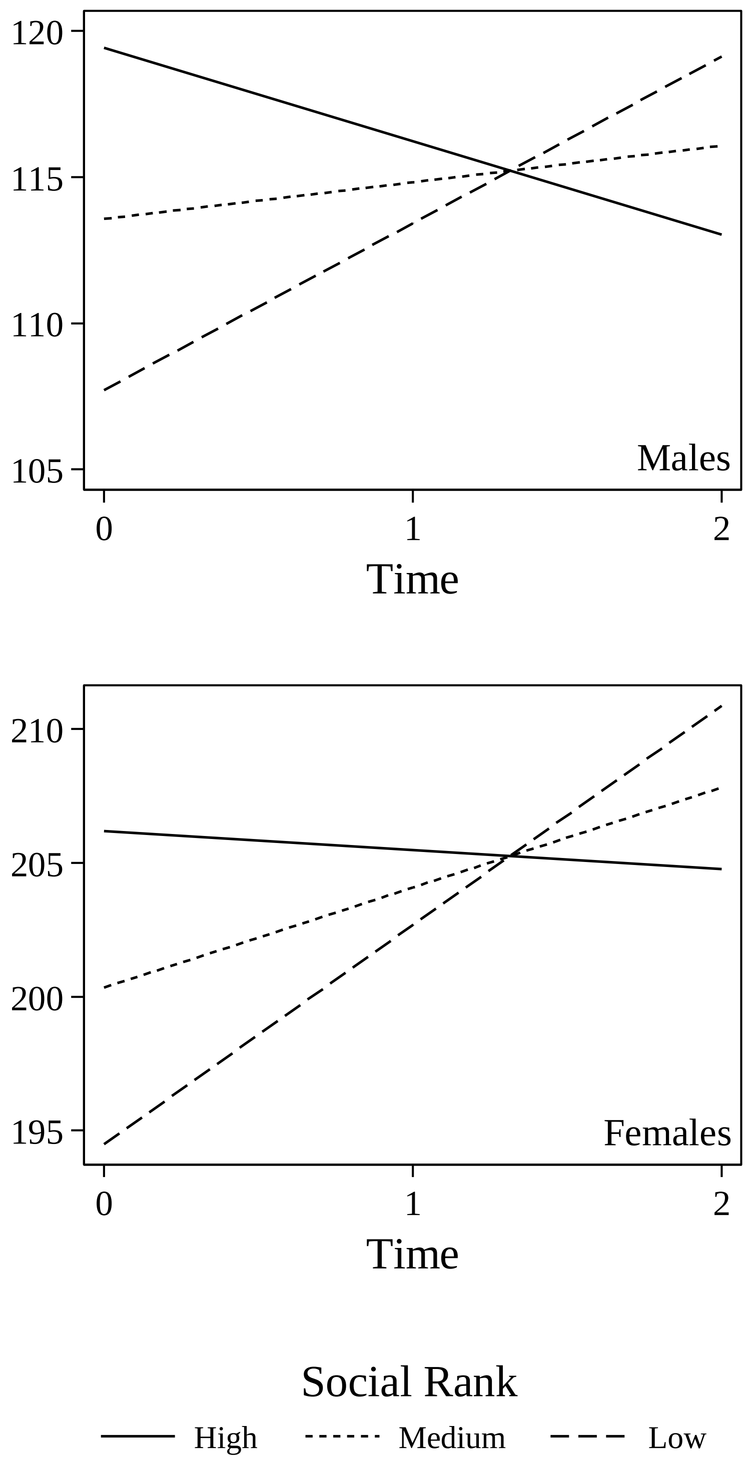
Figure 1. Fitted model of vocal pitch trajectory as a function of social rank. The dashed, dotted, and, solid line shows the pitch profile of individuals with a social rank score at the 10^{th}, 50^{th}, and 90^{th} percentile, respectively. In contrast to low-ranking individuals, whose pitch trajectory rises over time, the vocal trajectory of high-ranking individuals deepens over time.

Table 1.

Predictor variables	Regression Coefficient (*SE*)
Time	0.86
	(1.26)
Social Rank	5.85^{*}
	(2.43)
Time × Social Rank	-4.45^{***}
	(1.33)
Gender (female = 1)	86.78^{***}
	(3.61)
Time × Gender	2.48
	(1.89)
(Intercept)	114.08^{***}
	(2.41)
Observations	488
Clusters	173

* p<0.05, ** p<.01, *** p<.001

997 **Figure 1.**



998

Online Supplemental Material for

Listen, Follow Me: Dynamic Vocal Signals of Dominance Predict Emergent Social Rank in Humans

Joey T. Cheng[1*], Jessica L. Tracy[2], Simon Ho[2], Joseph Henrich[3]

[1] Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, Illinois 61820, USA
[2] Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia V6T 1Z4, Canada
[3] Department of Human Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138, USA

*To whom correspondence should be addressed. Email: jtcheng@illinois.edu

ONLINE SUPPLEMENTAL MATERIAL FOR DYNAMIC VOCAL SIGNALS OF DOMINANCE PREDICT EMERGENT SOCIAL RANK IN HUMANS

TABLE OF CONTENTS

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

These data derive from a larger project that was designed to examine the psychological underpinnings of social hierarchy. The current research complements this prior work, which focused on the behavioral strategies that promote rank (Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013) and the patterns of social attention in hierarchies (Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010), by addressing the novel question of how vocal changes influence the outcomes of rank contests.

Below we describe the calculation of our behavioral measure of social rank.

STUDY 1

ASSESSING BEHAVIORAL DECISION-MAKING IMPACT

Decision-making impact was measured by calculating the absolute difference between each participant's private ranking of each item on the Lost on the Moon task and his/her group's final ranking of that item, then summing across all 15 items and multiplying by -1, such that scores with a higher value (i.e., negative values closer to 0) reflect greater decisional impact. This scoring procedure can be represented as:

$$y_{ij} = -1\left(\sum_{k=1}^{15} \left|x_{ijk} - x_{jk}\right|\right)$$

where y_{ij} is the influence score of subject *i* from group *j*. x_{ijk} is subject *i*'s rating on item *k*. x_{jk} is group *j*'s rating on item *k*.

STUDY 2

EXPERIMENTAL STIMULI AND PARTICIPANT INSTRUCTIONS

In the master recording, a male research assistant read three statements into an Audio-Technica AT2020 microphone, positioned approximately 10cm away. Recordings were encoded in mono using Cool Edit Pro 2.1 at 44.1 kHz sampling rate and 16-bit quantization, and saved as three separate "wav" files. The statements are as follows:

> **Statement 1:** "Hi, I'm Joshua. I've worked here at PIQ as a sales and marketing director for six and a half years. I'm glad to be part of this team that will get to decide how the company's year-end bonus is to be allocated among our top 6 employees."
>
> **Statement 2:** "It's not that this second candidate, Luke Myers, isn't great. But he's only been with the company for four months. I think people who've worked here longer and made more contributions deserve a larger bonus."
>
> **Statement 3:** "Ok. I agree with what Jon said. We should give a larger share of the bonus to this fifth candidate, Michael Lawson. Look at his performance over the last year. He outshines everyone else in how well he's done."

As outlined in the main text, these segments were then raised and lowered in pitch, independent of other acoustic properties to create one version with a progressively deepening pitch trajectory across statements 1 to 3, and another version that contains a progressively rising pitch trajectory. All pitch manipulations were carried out in Praat phonetic analysis software using the pitch synchronous overlap add algorithm, and verified using the autocorrelation algorithm. The mean pitch of these segments in the two recordings is summarized in Table S1. Note that the two stimuli are identical in all respects other than pitch, including the average pitch across the entire duration of each recording.

TABLE S1. MEAN SUMMARY PITCH STATISTICS OF VOICE STIMULI

	Segment 1 (Hz)	Segment 2 (Hz)	Segment 3 (Hz)	Mean (Hz)
Deepening pitch recording	124.17	105.15	95.63	108.32
Rising pitch recording	94.21	105.15	124.41	107.92

Both pairs of vocal stimuli were presented to participants, who received the following written instructions at the outset of the experiment:

> "In this short survey, your task is to assist an employee who is trying to prepare for and rehearse a speech that he/she plans to give at the upcoming company meeting. In the meeting, this employee will be getting together with colleagues who, like him/her, are part of the company's year-end bonus allocation committee. Everyone present is the head of his/her specific division, and they will each be presenting their case for why the $50,000 year-end bonus should be awarded to one specific top-performing employee in their division who is working directly under their supervision. The final decision will be based on everyone's collective consensus, so they are each motivated to influence others into endorsing their division's candidate.
>
> Your task is to help one particular division head who, like the others, is trying to figure out how he/she should deliver his/her speech to most effectively influence the others. He/she has recorded two versions of how he/she could say what he/she wants to say to the committee. These two deliveries have identical content but differ in their speech styles.
>
> You are asked to listen carefully to both versions of this division head's speech rehearsal, and then indicate your perceptions of what the two versions convey.
>
> In both versions, he/she will give you a sample of the planned content at three different parts of his/her speech. The first segment features statements that he/she will make at the very beginning of the team's meeting, the second segment will be spoken at approximately 3 minutes into the meeting, and the third segment will be spoken at approximately 7 minutes into the meeting.
>
> On the next pages, you will listen to both versions, and each will be played in the form of a single video. Please play both videos, listen and watch very carefully, and then respond to a number of questions."

After listening to both pairs of recordings, participants were asked to respond to questions that would probe their perceptions of rank-seeking intentions and formidability. For each, they were asked to choose which of the pair of recordings they considered more characteristic. Specifically, the following questions were presented in randomized order:

Rank-Seeking Intentions:
"In which version did he appear to be trying to gain power?"
"In which version did he appear to be trying to be the leader of the committee?"
"In which version did he appear to be trying to control the committee's decision-making?"
"In which version did he appear to be trying to be assertive?"

Dominance:
"In which version did he appear to try to intimidate the other committee members?"
"In which version did he appear to try to be threatening?"

Prestige:
"In which version did he appear to try to gain the admiration of other committee members?"
"In which version did he appear to try to gain the respect of other committee members?"

SUPPLEMENTARY RESULTS AND ANALYSES

Here, we report additional results and analyses to supplement the primary findings presented in the main text.

STUDY 1

SAMPLE DEMOGRAPHIC INFORMATION AND SUMMARY STATISTICS

Below, we first present mean demographic information and summary statistics on the main variables. Second, we examine the robustness of our primary results to the inclusion of a number of additional controls by building on the baseline model presented in the main text. These controls include age, height, weight, group size, and postural expansiveness. Third, we explore the robustness of our conclusions from the growth model by using goodness of fit measures. Fourth, we present results of growth models using the three separate indices of rank—group member-rated rank, outside observer-rated rank, and behavioral decision-making impact—to supplement the results shown in the main text using the composite of these measures. Fifth, in another check of our conclusions, we present analyses that use a simple difference score approach to assess change in pitch. For these analyses, the pitch of each individual's first utterance was subtracted from the pitch of their third utterance, and these difference scores were regressed on social rank and controls. Finally, we examine the predictive capacity of absolute pitch levels (i.e., rather than change in pitch) on social rank outcomes. These analyses confirm that the observed differences between individuals in social rank cannot be reliably explained by absolute differences in pitch, and are instead a function of *changes* in vocal pitch.

TABLE S2. MEAN DESCRIPTIVE INFORMATION ON SAMPLE, PITCH PARAMETERS, AND SOCIAL RANK

	% of Sample	Mean	*SD*	Minimum	Maximum
Gender					
Men	54.01				
Women	45.99				
Ethnicity					
East Asian	46.52				
Caucasian	29.95				
South Asian	10.70				
Hispanic/Latino	1.60				
Mixed/Other	11.23				
Age		23.01	6.08	17	52
Group Size					
4-member group	14.97				
5-member group	26.74				
6-member group	54.55				
7-member group	3.74				
Pitch of 1st utterance (Hz)					
Men		114.08	17.15	79.46	171.43
Women		201.30	32.54	109.10	277.05
Pitch of 2nd utterance (Hz)					
Men		115.21	18.07	78.16	162.79
Women		206.61	28.10	161.11	274.29
Pitch of 3rd utterance (Hz)					
Men		114.91	15.19	81.10	163.44
Women		206.98	26.49	155.78	285.70
Social rank composite		-.003	.78	-2.28	2.00
Group member-rated rank		4.12	1.12	1.59	6.76
Outside observer-rated rank		2.80	1.00	1	5
Behavioral decision-making impact		-38.03	13.29	-84	-8
Perceived dominance		2.33	.81	1.19	5.43
Perceived prestige		4.93	.62	3.09	6.36

MISSING DATA

Although a total of 191 participants completed the study, the video-recording of one session of 4 participants (and as a result their speech samples) was unavailable due to equipment failure, resulting in a total of 187 participants for whom audio samples were available. However, vocal pitch parameters on all three utterances were indeterminable for 14 participants, who either did not speak in the group task or spoke too softly for speech to be recognized by the Praat phonetic analysis software. The analyses presented in the main text and below therefore contain data from 173 participants for whom any pitch parameters were available.

ALTERNATIVE MODEL SPECIFICATIONS WITH ADDITIONAL CONTROLS

In a series of individual growth models, we estimated the pitch of each utterance from time, social rank, and the interaction of time × social rank, as well as additional controls added sequentially, including gender, age, height and weight, group size, and postural expansiveness. Gender was dummy coded, with women = 1. Self-reported height was measured in feet, and weight in pounds. Group size was captured by two dummy variables, group size 5 and group size 6 or 7, for comparison to the reference group size of 4. Postural expansiveness was obtained from detailed coding of each participant's postural displays exhibited during the group task from video-recordings. To minimize capturing moment-to-moment random fluctuations in participants' displays and instead derive a more precise measure of expansiveness, one research assistant, who was blind to hypotheses, selected eight 20-second segments from throughout each group's entire interaction. These segments were chosen to feature moments of key decisions and discussion in the group. Two other trained research assistants viewed these eight segments and independently coded the extent to which each participant "occupied room with the body", "held a wide posture", and "extended arms out from the body", on a 6-point Likert scale (0 = *Not at all present*; 5 = *Extreme intensity*). For each item, we summed the rating across all eight segments and then averaged across the two coders (i.e., aggregating across the 8 ratings × 2 coders = 16 observations for each participant). Finally, each individual's score on the three items were averaged to form an overall measure of postural expansiveness (α= .91). All continuous predictors—age, height, weight, and postural expansiveness—were grand mean centered in the model.

To examine the effects of these controls, compare Models 1 (baseline, no controls) and 2 (with gender as control, reported in the main text) with Models 3-6 (with additional controls) in Table S3 below. The coefficients on the time × social rank interaction term in Models 2-5, ranging from -4.80 to -4.32, remain stable from model to model and comparable to that of Model 2 (estimated to be -4.45), as well as statistically significant. In all models, apart from a significant gender main effect, none of the controls were significant at conventional levels. Model 2 showed the lowest BIC, indicating best fit adjusted for both model and sample size, so we reported results from this specification in the main text. Overall, these results converge to indicate that the diverging pitch trajectories found for low- and high-ranking individuals are unlikely to be driven by differences in age, gender, body size, or by dynamics of the local group, such as the number of group members present or one's physical expansiveness during the interaction. Taken together, results from all six models converge to indicate that, independent of these demographic and contextual factors, each 1-point increase in social rank was associated with a pitch trajectory that lowered approximately 4-5 Hz over each utterance.

TABLE S3. HIERARCHICAL LINEAR MODELS PREDICTING VOCAL PITCH FROM THE MAIN AND INTERACTIVE EFFECTS OF TIME AND SOCIAL RANK COMPOSITE, AS WELL AS CONTROL VARIABLES: AGE, GENDER (1 = WOMEN), HEIGHT, WEIGHT, GROUP SIZE DUMMIES, AND POSTURAL EXPANSIVENESS, AND THEIR INTERACTION WITH TIME. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

Predictor variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Time	1.82†	0.86	0.89	1.45	3.60	4.12
	(0.96)	(1.26)	(1.26)	(1.53)	(2.66)	(2.71)
Social Rank	9.93*	5.85*	5.84*	7.00**	6.93**	8.88***
	(5.02)	(2.43)	(2.42)	(2.49)	(2.47)	(2.59)
Time × Social Rank	-4.44***	-4.45***	-4.45***	-4.39**	-4.32**	-4.80***
	(1.34)	(1.33)	(1.32)	(1.37)	(1.37)	(1.43)
Gender (Female=1)		86.78***	86.96***	80.45***	80.70***	82.54***
		(3.61)	(3.62)	(5.05)	(5.08)	(4.99)
Time × Gender		2.48	2.42	1.14	1.28	0.66
		(1.89)	(1.89)	(2.67)	(2.71)	(2.70)
Age			-0.19	-0.04	-0.12	-0.34
			(0.30)	(0.31)	(0.32)	(0.32)
Time × Age			0.07	0.06	0.09	0.16
			(0.15)	(0.16)	(0.17)	(0.17)
Height (feet)				-8.96	-8.75	-7.84
				(7.59)	(7.52)	(7.35)
Weight (lbs)				-0.06	-0.04	-0.02
				(0.07)	(0.07)	(0.07)
Time × Height				-4.40	-4.43	-4.80
				(3.98)	(3.97)	(3.95)
Time × Weight				0.02	0.01	0.01
				(0.03)	(0.04)	(0.04)
Group Size 5 (dummy)					8.79	8.44
					(5.98)	(6.04)
Group Size 6 or 7 (dummy)					9.38†	8.61†
					(5.05)	(5.09)
Time × Group Size 5					-3.27	-3.47
					(3.13)	(3.22)
Time × Group Size 6 or 7					-2.36	-2.35
					(2.65)	(2.71)

Postural Expansiveness						-0.26
						(0.46)
Time × Postural Expansiveness						-0.01
						(0.25)
(Intercept)	152.31***	114.08***	113.99***	117.00***	109.14***	108.24***
	(3.75)	(2.41)	(2.41)	(2.89)	(5.08)	(5.09)
AIC	4678.64	4367.45	4371.04	4369.75	4373.87	4359.55
BIC	4712.16	4409.35	4421.32	4436.79	4457.68	4451.69
Observations	488	488	488	488	488	488
Clusters	173	173	173	173	173	172

† p<0.10, * p<0.05, ** p<.01, *** p<.001

MODEL FIT INFORMATION

To further support the analyses reported in Table S3 and the main text, which tested the significance of the time × social rank cross-level interaction using null hypothesis testing (i.e., testing whether the ratio of the coefficient to the standard error differs from zero; Raudenbush & Bryk, 2002), we examine model fit here. We compared the Akaike information criterion (AIC) and Bayesian information criterion (BIC) of models that include the cross-level time × social rank interaction for each of the five specifications, presented above in Table S3, with their corresponding nested models in which the interaction term is omitted.

To examine the effects of the time × social rank interaction, compare the AIC and BIC of Models 1 to 6 reported in Table S3 above with the fit indices of Models 1’ to 6’ in Table S4 below. Across all six specifications, the AIC and BIC are lower in the original model in Table S3 above with the interaction term, than in the nested model in Table S4 without the interaction term, indicating better fit of the models with the interaction term, across the board. The differences in AIC range from 9.12 to 8.17 and the differences in BIC range from 4.94 to 3.62, all of which exceed the threshold value, 2, for positive evidence indicating a difference in fit (Raftery, 1995). These results reconfirm the predictive importance of the time × social rank interaction, and by implication the conclusion that higher social rank is associated with a deepening pitch profile over time.

TABLE S4. HIERARCHICAL LINEAR MODELS PREDICTING VOCAL PITCH FROM THE MAIN EFFECTS OF TIME AND SOCIAL RANK COMPOSITE, AS WELL AS CONTROL VARIABLES INCLUDING AGE, GENDER, HEIGHT, WEIGHT, GROUP SIZE DUMMIES, AND POSTURAL EXPANSIVENESS, AND THEIR INTERACTION WITH TIME. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES. MODELS 1' TO 6' HERE, WHICH DO NOT INCLUDE THE TIME × SOCIAL RANK INTERACTION TERM, CORRESPOND TO MODELS 1 TO 6 IN TABLE S3 ABOVE THAT CONTAIN THE INTERACTION. FOR BOTH FIT INDICES (AIC AND BIC), LOWER VALUES INDICATE BETTER FIT.

	Model 1'	Model 2'	Model 3'	Model 4'	Model 5'	Model 6'
Time	1.05	0.38	0.41	1.64	3.88	4.93†
	(0.95)	(1.28)	(1.28)	(1.55)	(2.70)	(2.74)
Social Rank	7.02	1.13	1.13	2.21	2.28	3.67†
	(4.94)	(1.98)	(1.98)	(2.00)	(1.99)	(2.09)
Gender (Female=1)		87.29***	87.48***	82.59***	82.64***	84.16***
		(3.64)	(3.65)	(5.04)	(5.08)	(5.01)
Time × Gender		2.05	1.99	-0.62	-0.33	-0.94
		(1.92)	(1.92)	(2.66)	(2.71)	(2.70)
Age			-0.20	-0.10	-0.18	-0.38
			(0.30)	(0.32)	(0.32)	(0.32)
Time × Age			0.06	0.10	0.13	0.20
			(0.16)	(0.17)	(0.17)	(0.17)
Height (feet)				-7.43	-7.31	-6.56
				(7.64)	(7.56)	(7.41)
Weight (lbs)				-0.04	-0.01	0.01
				(0.07)	(0.07)	(0.07)
Time × Height				-5.69	-5.66	-6.12
				(4.03)	(4.02)	(3.99)
Time × Weight				0.00	-0.01	-0.02
				(0.03)	(0.04)	(0.04)
Group Size 5 (dummy)					9.48	10.01†

					(6.02)	(6.07)
Group Size 6 or 7 (dummy)					9.48†	9.49†
					(5.09)	(5.13)
Time × Group Size 5					-3.81	-4.80
					(3.18)	(3.25)
Time × Group Size 6 or 7					-2.37	-3.01
					(2.69)	(2.75)
Postural Expansiveness						0.03
						(0.45)
Time × Postural Expansiveness						-0.25
						(0.24)
(Intercept)	152.59***	114.31***	114.21***	116.45***	108.43***	107.08***
	(3.76)	(2.43)	(2.43)	(2.91)	(5.11)	(5.12)
AIC	4687.42	4376.52	4380.09	4377.92	4381.68	4368.67
BIC	4716.75	4414.23	4426.19	4440.77	4461.30	4456.63
Observations	488	488	488	488	488	487
Clusters	173	173	173	173	173	172

† $p<0.10$, * $p<0.05$, ** $p<.01$, *** $p<.001$

DO MEN AND WOMEN SHOW DIFFERENT RANK-RELATED PITCH TRAJECTORIES?

Do men and women use different kinds of vocal pitch-based strategies to signal and compete for social rank? Cultural explanations of gender differences, such as the notion of "gender code", predict a sex difference in this respect (Cartei, Cowles, & Reby, 2012); specifically, cultural beliefs, practices, and institutions expect and reward the demonstration of masculinity in men and femininity in women, and also link masculinity (and the display of masculine traits and characteristics) with success and competition. These strong cultural associations may result in men and women conforming to the gender code by systematically manipulating their vocal pitch to be consistent with gender norms.

This account predicts that whereas men may readily acquire—by trial and error, imitation, or learning—the propensity to flaunt signs of masculinity, including deepening their voice in social situations, women may raise their pitch or show little systematic vocal modulation.

To test this account, we examined whether the association between pitch trajectories and social rank is moderated by the signaler's gender. It is important to note, however, that these analyses should be interpreted with caution, due to the limited statistical power afforded by our relatively small sample of men (n = 96) and women (n = 77) for precisely estimating within-gender effects in individual growth models. Statistical power and sample size concerns in estimating cross-level interactions in multilevel models are particularly complex and jointly influenced by interactive relationships among multiple factors, many of which are not relevant in single-level analyses (Mathieu, Aguinis, Culpepper, & Chen, 2012; Scherbaum & Ferreter, 2009; Snijders, 2005; Snijders & Bosker, 1993, 1994). Simulation studies (Bassiri, 1988; van der Leeden & Busing, 1994) suggest that whereas cross-level interactions generally require 30 observations at both the higher and lower level, fewer observations at either level leads to a rapid decline of power and precision for detecting cross-level interactions, such that with only five observations at the lower level, 150 upper level units are needed (for a review, see Kreft & de Leeuw, 1998). This suggests that, for our growth model, which has 3 repeated observations of pitch per individual, at least 150 subjects are needed in order to detect and estimate the cross-level time × social rank interaction—a criterion that is met by our combined sample size of 173 participants, but not by either sample of men or women alone. Thus, any estimated effects of within-gender relationships may also be biased (Snijders & Bosker, 1993).

Nevertheless, results from a growth model that accounts for the possibility of gender moderation indicate no qualitative difference between men and women in their pitch strategies. Although the growth model revealed a significant time × social rank × gender interaction in predicting vocal pitch (b = -4.67, SE = 2.19, p = .033, .95CI[-8.97, -.37]; see Table S5 below for full model results), descriptively the pitch profiles for high- and low-ranking individuals are similar across both genders. Specifically, the model estimates a 3.40 Hz *increase* per utterance for low-ranking men, and a 11.97 Hz increase per utterance for low-ranking women, with a significantly steeper trajectory for women than for men (p = .012). Moreover, the model estimates a 1.74 Hz *decrease* per utterance for high-ranking men and a 2.61 Hz *decrease* for high-ranking women, and these two trajectories do not differ significantly (p = .722). Thus, these analyses indicate that for both men and women, a lowering pitch profile is associated with greater social rank, whereas a raising pitch profile is associated with lower social rank. Furthermore, the only gender difference that did emerge suggests that low-ranking women raise their pitch to a greater degree than their low-ranking male counterparts, whereas high-ranking men and women lower their pitch to a similar degree.

In sum, although these results are in need of future replication with studies using larger samples, these analyses offer very preliminary support for the notion that pitch change signals operate similarly across gender. These results are consistent with prior work indicating that lower voices function to increase dominance in women (as well as men), including findings that lower pitched female voices are perceived as more dominant and successful in acquiring resources (e.g., Apicella & Feinberg, 2009; Borkowska & Pawlowski, 2011; Jones, Feinberg, DeBruine, Little, & Vukovic, 2010), and that speaking in a lower pitch (following experimental instructions) enhances women's subjective reports of power and dominance, as it does in men (Stel, Dijk, Smith, Dijk, & Djalal, 2012). The finding that low-ranking women raise their voice to a greater magnitude than similarly ranked men may be due to anatomical differences and constraints that equip women with a wider possible range of pitch (e.g., Cartei et al., 2012).

TABLE S5. HIERARCHICAL LINEAR MODELS PREDICTING VOCAL PITCH FROM THE MAIN EFFECTS OF TIME, SOCIAL RANK COMPOSITE, GENDER, THE INTERACTION AMONG THESE VARIABLES. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES. THIS MODEL CORRESPONDS TO MODEL 2 IN TABLE S3 ABOVE BUT WITH THE ADDITION OF THE TIME × SOCIAL RANK × GENDER INTERACTION TERM.

Predictor variables	Regression Coefficient (*SE*)
Time	0.61
	(1.26)
Social Rank	5.86*
	(2.43)
Gender (female = 1)	86.73***
	(3.62)
Time × Social Rank	-2.57
	(1.59)
Time × Gender	3.39
	(1.92)
Time × Social Rank × Gender	-4.67*
	(2.19)
(Intercept)	114.10***
	(2.41)
Observations	488
Clusters	173

* p<0.05, ** p<.01, *** p<.001

ROBUSTNESS CHECK USING THE THREE SEPARATE INDICES OF SOCIAL RANK

To check the robustness of our findings to each of the three specific indices that, in the main text, were aggregated into a social rank composite, and paralleling the results reported above, we examined two specifications, one of which controls for gender (the best fitting specification shown in the main text and Model 2 in Table S3) and the other of which includes the full set of controls (Model 6 in Table S3).

Tables S6 and S7 display these two sets of individual growth models. Coefficients on the time × rank index term in the models using outside observer-rated rank and behavioral decision-making impact were sizable, negative, and significant, but were not well estimated for group member-rated rank (*ps* = .051 and .074, respectively). Nevertheless, these results are generally consistent with each other and with findings based on the rank composite, indicating that our conclusions are robust across diverse approaches to measuring social rank. Note that the magnitude of the coefficients on social rank cannot be directly compared because units vary depending on the specific rank index.

TABLE S6. HIERARCHICAL LINEAR MODELS PREDICTING VOCAL PITCH FROM THE MAIN AND INTERACTIVE EFFECTS OF TIME AND SOCIAL RANK INDEX (WHICH DIFFERS IN EACH MODEL), AS WELL AS THE CONTROLS GENDER AND ITS INTERACTION WITH TIME. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

	Social Rank Index		
	Group member-rated rank	Outside observer-rated rank	Behavioral decision-making impact
Time	0.67	1.11	0.28
	(1.28)	(1.28)	(1.27)
Social Rank	2.42	3.77†	0.30*
	(1.71)	(1.94)	(0.14)
Time × Social Rank	-1.79†	-3.30**	-0.19**
	(0.92)	(1.07)	(0.07)
Gender (Female=1)	87.06***	87.41***	86.64***
	(3.64)	(3.62)	(3.65)
Time × Gender	2.26	2.03	2.59
	(1.91)	(1.89)	(1.91)
(Intercept)	114.18***	113.85***	114.56***

	(2.43)	(2.43)	(2.43)
Observations	488	488	485
Clusters	173	173	172

† p<0.10, * p<0.05, ** p<.01, *** p<.001

TABLE S7. HIERARCHICAL LINEAR MODELS PREDICTING VOCAL PITCH FROM THE MAIN AND INTERACTIVE EFFECTS OF TIME AND SOCIAL RANK INDEX (WHICH DIFFERS IN EACH MODEL), AS WELL CONTROL VARIABLES AGE, GENDER, HEIGHT, WEIGHT, GROUP SIZE DUMMIES, AND POSTURAL EXPANSIVENESS, AND THEIR INTERACTION WITH TIME. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

	Measure of Social Rank		
Predictor variables	Group member-rated rank	Outside observer-rated rank	Behavioral decision-making impact
Time	4.76†	5.17†	3.77
	(2.74)	(2.72)	(2.77)
Social Rank	3.93*	5.62**	0.37**
	(1.80)	(2.06)	(0.14)
Time × Social Rank	-1.74†	-3.53**	-0.20**
	(0.97)	(1.15)	(0.07)
Age	-0.42	-0.36	-0.32
	(0.32)	(0.32)	(0.32)
Gender (Female=1)	83.54***	84.18***	83.59***
	(5.08)	(5.00)	(5.06)
Time × Age	0.20	0.17	0.15
	(0.17)	(0.17)	(0.17)
Time × Gender	-0.13	-0.15	0.07
	(2.73)	(2.68)	(2.72)
Height (feet)	-7.93	-6.74	-5.86
	(7.52)	(7.42)	(7.46)
Weight (lbs)	0.01	-0.00	-0.00
	(0.07)	(0.07)	(0.07)
Time × Height	-5.07	-5.32	-5.88
	(4.02)	(3.96)	(3.98)
Time × Weight	-0.01	0.00	-0.00
	(0.04)	(0.04)	(0.04)
Group Size 5 (dummy)	9.89	10.39†	8.67
	(6.12)	(6.06)	(6.15)
Group Size 6 or 7 (dummy)	9.82†	10.01†	7.96
	(5.16)	(5.13)	(5.20)
Time × Group Size 5	-4.33	-4.37	-3.60
	(3.25)	(3.22)	(3.26)
Time × Group Size 6 or 7	-2.97	-3.06	-2.03
	(2.74)	(2.72)	(2.76)
Postural Expansiveness	-0.04	-0.21	0.11
	(0.46)	(0.47)	(0.45)

Time × Postural Expansiveness	-0.15	0.00	-0.20
	(0.25)	(0.25)	(0.24)
(Intercept)	107.00***	106.29***	108.71***
	(5.16)	(5.12)	(5.23)
Observations	487	487	484
Clusters	172	172	171

† p<0.10, * p<0.05, ** p<.01, *** p<.001

ASSESSING CHANGE IN PITCH WITH DIFFERENCE SCORES

As an alternative check of our results, we conducted analyses using the simple difference between the pitch of the third utterance and first utterance—the two available points that differed maximally in time—to assess change (Rogosa & Willett, 1983). We calculated a measure of change in pitch for each individual by subtracting the pitch of the first utterance from the pitch of the third utterance. If a group member raised her pitch over time, her pitch change score would take on a positive value, whereas if she lowered her pitch, her change score would take on a negative value.

Table S8 shows a series of linear regression models regressing pitch change on the social rank composite (grand mean centered), and sequentially added controls. Independent of the controls, none of which were significant, the coefficients on social rank were negative and significant at conventional levels across all seven models. A 1-point increase in social rank was associated with a decrease in pitch ranging from roughly -8.1 to -8.7 Hz. This corresponds to approximately a 6% change in fundamental frequency on average for each 1-point increase in social rank, a magnitude that prior perceptual studies have revealed to meet the discrimination threshold for humans, which for this frequency range generally ranges from 2 to 7% conditional on the specific type of voice material (see Ives, Smith, & Patterson, 2005; Pisanski & Rendall, 2011; Puts, Hodges, Cárdenas, & Gaulin, 2007; Sinnott, Owren, & Petersen, 1987; Smith, Patterson, Turner, Kawahara, & Irino, 2005), and is thus perceptually relevant. Note, however, that this measured magnitude of change reflects pitch modulations observed over the initial vocalizations of the interaction. Given the brief time frame over which change is examined, our measure likely underestimates the true strength of the pitch change signals deployed. The actual salience and strength of pitch changes in rank contests, and whether these signals can be perceptually discriminated, is an important question for future work (see the General Discussion in the main text for further discussion). Nevertheless, these models estimate that, independent of the controls, the average low-ranking individual who scores at the 10th percentile on social rank raises his/her pitch by approximately 12.13 to 12.82 Hz from the first to third utterance. In contrast, the average high-ranking individual who scores at the 90th percentile on social rank *lowers* his/her pitch by approximately 4.53 to 4.96 Hz from the first to third utterance.

TABLE S8. ORDINARY LEAST SQUARE REGRESSIONS FOR SIMPLE DIFFERENCE BETWEEN PITCH AT THE THIRD AND FIRST UTTERANCES, WITH SOCIAL RANK. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

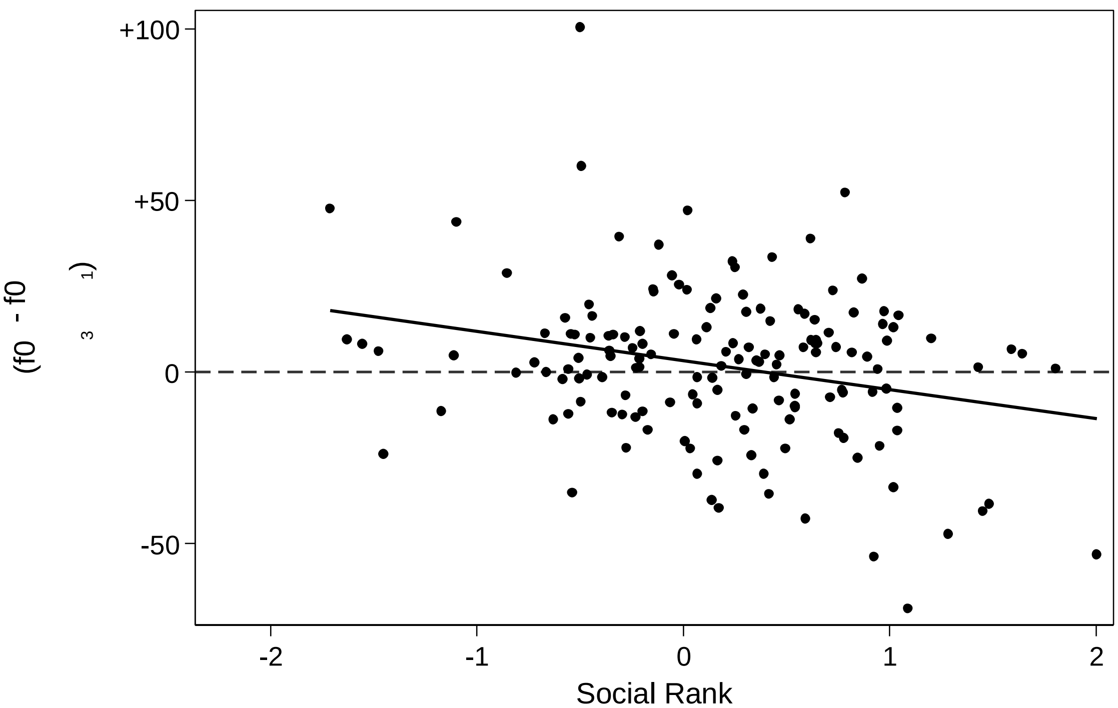
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Social Rank (Composite)	-8.48**	-8.63***	-8.65***	-8.48**	-8.31**	-8.11**	-8.11**
	(2.56)	(2.57)	(2.57)	(2.65)	(2.66)	(2.78)	(2.78)
Gender (Female=1)		2.80	2.58	-0.01	0.51	0.46	0.46
		(3.56)	(3.58)	(5.09)	(5.17)	(5.19)	(5.19)
Age			0.22	0.20	0.26	0.26	0.26
			(0.29)	(0.31)	(0.31)	(0.32)	(0.32)
Height (feet)				-8.50	-8.22	-8.32	-8.32
				(7.51)	(7.52)	(7.55)	(7.55)

Weight (lbs)				0.04	0.02	0.02	0.02
				(0.06)	(0.07)	(0.07)	(0.07)
Group Size 5 (dummy)					-8.18	-8.58	-8.58
					(5.90)	(6.13)	(6.13)
Group Size 6 or 7 (dummy)					-5.63	-5.95	-5.95
					(4.99)	(5.16)	(5.16)
Postural Expansiveness						-0.12	-0.12
						(0.47)	(0.47)
(Intercept)	3.38†	2.17	2.26	3.39	8.50†	8.80†	8.80†
	(1.83)	(2.39)	(2.40)	(2.90)	(5.01)	(5.16)	(5.16)
AIC	1389.27	1390.64	1392.06	1394.68	1396.60	1398.53	1398.53
BIC	1395.34	1399.75	1404.21	1412.90	1420.90	1425.86	1425.86
N	154	154	154	154	154	154	154

† p<0.10, * p<0.05, ** p<.01, *** p<.001

Figure S1 plots change in pitch—which, as can be seen, shows sizable variation between individuals—as a function of social rank. Confirming the results from the individual growth models reported in the main text and above, high-ranking individuals tended to deepen their pitch over the course of the interaction, whereas low ranking individuals raised their pitch.

FIGURE S1. SCATTER PLOT (WITH BEST-FITTING REGRESSION LINE) OF CHANGE IN VOCAL PITCH AS A FUNCTION OF SOCIAL RANK. CHANGE IN PITCH WAS COMPUTED BY SUBTRACTING THE PITCH OF AN INDIVIDUAL'S FIRST UTTERANCE FROM THE PITCH OF HIS OR HER THIRD UTTERANCE. THE DOTTED LINE SEPARATES INDIVIDUALS WHO LOWERED THEIR PITCH (BELOW THIS LINE) FROM THOSE WHO RAISED THEIR PITCH (ABOVE THIS LINE).



Note. On average, each 1-point incase in social rank was associated with a 8.48 Hz drop in vocal pitch. This negative association remains significant when the participant who showed the greatest pitch change (+100.42 Hz) is removed from the analysis; the model still estimates a 7.6 Hz drop in pitch for each unit of social rank (p = .002).

ABSOLUTE AND HABITUAL PITCH AND SOCIAL RANK

To examine whether mean pitch levels—as opposed to *changes* in pitch—are related to rank, we correlated the rank composite with the pitch parameter of each utterance and with mean pitch across all three utterances. With one exception, all correlations were non-significant at conventional levels: $f0_1$ (r = .05, p = .630 for males), $f0_2$ (r = .10, p = .346 for males; r = .003, p = .981 for females), $f0_3$ (r = -.02, p = .881 for males; r = -.17, p = .179 for females), and $f0_{mean}$ (reflecting habitual pitch; r = -.01, p = .951 for males; r = .18, p = .13 for females). However, the first utterance for females was significant, though in the opposite direction than expected (r = .28, p = .013). These results, combined with findings above, indicate that little of the variation between individuals in rank is explained by higher or lower mean momentary pitch levels or habitual pitch. Instead, rank is more directly linked to dynamics shifts in pitch over time.

PITCH MODULATIONS ARE RELATED TO DOMINANCE-BASED RANK, BUT NOT PRESTIGE-BASED RANK

Our analyses indicated that, as expected, individuals whose speech was characterized by a progressively deepening pitch profile attained greater dominance-based rank. Also as expected, prestige-based rank was not reliably associated with pitch change. In a growth model, pitch was modeled as a function of time of utterance, peers' ratings of dominance, peers' ratings of prestige, and the cross-level interaction between time and each of these rank-attainment strategies. The full series of models with a variety of controls are displayed in Table S9.

In the specification with only gender as control (Model 2; which yielded the lowest model BIC among all specifications and was reported in the main text), the time × dominance interaction was significant and indicates a 2.84 Hz decrease in pitch per utterance for each 1-point increase in dominance. In contrast, the time × prestige interaction was not significant. In the other models with controls (Models 3-6) and without controls (Model 1), the coefficient on time × dominance remained negative, varying from -2.80 to -3.04. As above, none of the time × prestige coefficients were significant. Notably, the pattern of results is robust across all of these alternative specifications.

TABLE S9. HIERARCHICAL LINEAR MODELS PREDICTING VOCAL PITCH FROM THE MAIN AND INTERACTIVE EFFECTS OF TIME, DOMINANCE, AND PRESTIGE, AS WELL AS CONTROL VARIABLES AGE, GENDER, HEIGHT, WEIGHT, GROUP SIZE DUMMIES, AND POSTURAL EXPANSIVENESS, AND THEIR INTERACTION WITH TIME. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Time	1.54	0.94	1.03	1.84	3.73	4.49
	(0.96)	(1.29)	(1.29)	(1.55)	(2.70)	(2.75)
Dominance	-0.41	3.81†	4.23†	4.83*	4.53*	5.34*
	(4.61)	(2.23)	(2.27)	(2.28)	(2.27)	(2.33)
Prestige	10.12	1.59	1.19	2.13	2.93	4.19
	(6.41)	(3.11)	(3.13)	(3.18)	(3.21)	(3.19)
Time × Dominance	-2.80*	-2.84*	-3.04*	-2.86*	-2.83*	-2.93*
	(1.16)	(1.17)	(1.20)	(1.21)	(1.22)	(1.26)
Time × Prestige	-1.38	-1.07	-0.89	-0.56	-0.55	-0.93
	(1.64)	(1.65)	(1.66)	(1.69)	(1.73)	(1.72)
Gender (Female=1)		87.74***	88.15***	82.51***	82.28***	84.10***
		(3.67)	(3.68)	(5.10)	(5.12)	(5.07)
Time × Gender		1.69	1.50	-0.35	0.02	-0.57
		(1.92)	(1.93)	(2.69)	(2.72)	(2.71)
Age			-0.30	-0.19	-0.26	-0.48
			(0.31)	(0.32)	(0.32)	(0.33)
Time × Age			0.14	0.15	0.18	0.25
			(0.16)	(0.17)	(0.17)	(0.17)
Height (feet)				-9.23	-9.42	-8.99
				(7.81)	(7.74)	(7.60)
Weight (lbs)				-0.04	-0.01	0.02
				(0.07)	(0.07)	(0.07)
Time × Height				-4.58	-4.54	-4.80
				(4.09)	(4.09)	(4.07)
Time × Weight				0.01	-0.00	-0.01
				(0.03)	(0.04)	(0.04)
Group Size 5 (dummy)					9.54	9.93
					(6.03)	(6.10)
Group Size 6 or 7 (dummy)					9.33†	9.20†
					(5.18)	(5.25)
Time × Group Size 5					-3.64	-4.17
					(3.16)	(3.24)
Time × Group Size 6 or 7					-1.90	-2.21
					(2.73)	(2.81)
Postural Expansiveness						-0.09
						(0.46)

Time × Postural Expansiveness						-0.09
						(0.25)
(Intercept)	152.58***	113.87***	113.67***	116.26***	108.41***	107.02***
	(3.78)	(2.44)	(2.44)	(2.91)	(5.13)	(5.16)
AIC	4686.68	4376.78	4379.73	4378.54	4382.19	4369.61
BIC	4728.59	4427.06	4438.40	4453.97	4474.38	4470.13
N	488	488	488	488	488	487

† p<0.10, * p<0.05, ** p<.01, *** p<.001

Moreover, additional analyses indicate that the effect of a deepening pitch profile in predicting greater dominance-based rank did not depend upon obtaining a specific level of prestige. When a three-way interaction term (time × dominance × prestige) was added to the same baseline model as above, which used only gender as control (Model 2), the three-way interaction was not statistically significant (b = 1.44, SE = 1.38, p = .295, .95CI[-1.25, 4.14]). Moreover, replicating our previous analysis, the time × dominance interaction term remained a significant negative predictor of vocal pitch (b = -2.63, SE = 1.19, p = .027, .95CI[-4.96, -.29]), whereas the time × prestige term had a null effect (b = -1.35, SE = 1.66, p = .418, .95CI[-4.61, 1.92]. These results offer further support to the notion that the link between pitch modulations and emergent social rank is underpinned by increased perceptions of dominance, but not prestige.

As a robustness check, we next turned to the alternative operationalization of pitch change using simple difference scores. We estimated a series of linear regression models in which pitch change—calculated by subtracting pitch of the first utterance from the third utterance—was regressed on dominance and prestige (both grand mean centered), along with, in some models, additional controls. While these models in Table S10 below show no effects of prestige or any of the controls, the coefficient on dominance remained significant and negative across all specifications, except in Model 6 where the effect was marginally significant (p = .053) when the full set of controls was included. The coefficients on dominance indicate that a 1-point increase in dominance predicts a decrease in vocal pitch between 4.79 Hz and 5.54 Hz, independent of the controls. These models estimate that, on average, as dominance increases from the 10th percentile to the 90th percentile, the estimated pitch trajectory changes from a rising profile, marked by an approximate increase between 6.49 and 7.22 Hz, to a lowering profile, in which pitch *lowers* by approximately -1.61 to -2.14 Hz.

TABLE S10. ORDINARY LEAST SQUARE REGRESSIONS PREDICTING PITCH CHANGE, OPERATIONALIZED AS PITCH ON THE THIRD UTTERANCE MINUS PITCH ON THE FIRST UTTERANCE, FROM DOMINANCE AND PRESTIGE. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Dominance	-5.10*	-5.02*	-5.54*	-5.16*	-5.14*	-4.79†
	(2.24)	(2.26)	(2.30)	(2.34)	(2.35)	(2.45)
Prestige	-2.11	-2.18	-1.80	-1.08	-1.22	-1.13
	(3.15)	(3.17)	(3.18)	(3.25)	(3.33)	(3.35)
Gender (Female=1)		1.24	0.76	-2.97	-1.98	-1.98
		(3.67)	(3.69)	(5.16)	(5.23)	(5.24)
Age			0.34	0.37	0.44	0.43
			(0.31)	(0.32)	(0.33)	(0.33)
Height (feet)				-8.89	-8.40	-8.65
				(7.81)	(7.81)	(7.84)
Weight (lbs)				0.01	-0.01	-0.01
				(0.07)	(0.07)	(0.07)
Group Size 5 (dummy)					-9.05	-9.84
					(6.02)	(6.22)
Group Size 6 or 7 (dummy)					-4.87	-5.58
					(5.21)	(5.40)
Postural Expansiveness						-0.25
						(0.48)
(Intercept)	2.67	2.12	2.35	3.97	8.70†	9.31†
	(1.85)	(2.47)	(2.48)	(2.97)	(5.15)	(5.29)
AIC	1396.75	1398.63	1399.33	1401.93	1403.54	1405.25
BIC	1405.86	1410.78	1414.52	1423.19	1430.87	1435.62
N	154	154	154	154	154	154

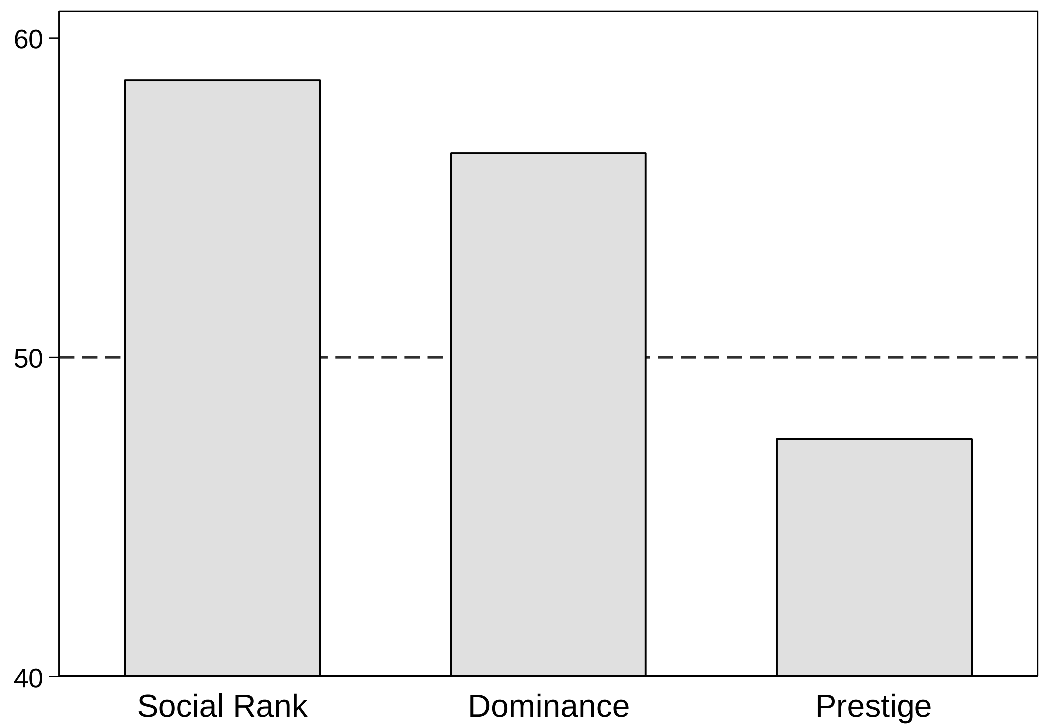
† p<0.10, * p<0.05, ** p<.01, *** p<.001

STUDY 2

FORCED CHOICE SUMMARY DATA

Figure S2 summarizes the data from Study 2. As can be seen, the proportion of participants who matched the social rank composite and the domiance composite to the deepening pitch recording (compared to the alternative rising pitch recording) was significantly greater than would be expected by chance (i.e., 50%).

FIGURE S2. THE PROPORTION OF PARTICIPANTS WHO SELECTED THE DEEPENING PITCH RECORDING FOR EACH TRAIT. THE DOTTED LINE INDICATES CHANCE



FORCED CHOICE RESULTS WITH CONTROLS

To test whether these effects are robust to the inclusion of subject-level controls, we regressed the summative measure of each participant's average tendency to select the deepening voice on gender, age, method of recruitment (student or Mechanical Turk worker), and stimuli presentation order. With the exception of age, which was the only continuous predictor, all dichotomous predictors were represented by effects coding. Table S10 shows three regression models using these controls to predict the average tendency to select the deepening voice for communicating greater rank-seeking intention, dominance, and prestige. With effects coding, the intercept in each model summarizes the model predicted unweighted grand mean preference for a deepening voice (coded ‘0’ for each individual item) across all participants, accounting for the effects of the full set of controls. An intercept of .50 is the expected value if choice is completely random and unaffected by the pitch manipulation. The coefficient on the predictors gives the deviation of the choice ratio for the group coded ‘+1’ and the grand mean.

In both the models predicting social rank and dominance, the intercept values of .32 and .21 deviate substantially from .50 and are skewed toward 0, indicating an overall tendency to choose the deepening voice recording as more descriptive of these traits. In contrast, the intercept in the model predicting perceptions of prestige is .53, which suggests a slight bias toward choosing the *rising* pitch recording.

Wald test was used to investigate whether the intercept in each model differs from .50, which indicates random choice. As expected, the intercept in the dominance model differed significantly from .50 [$F(1, 269) = 4.48$, $p = .035$], whereas the intercept in the prestige model did not differ from .50 [$F(1, 269) = .07$, $p = .795$]. The dominance effect was thus robust to the inclusion of these subject-level controls. The intercept in the social rank model, however, was only marginally significant at conventional levels [$F(1, 269) = 2.50$, $p = .115$]

To further explore the effects of these controls on the social rank composite, we re-ran the same regression models but sequentially added gender, age, stimuli presentation order, and method of recruitment to a baseline model (with no predictors), and again in each case examined the estimated intercept, which provides an unweighted grand mean of voice preference across all subjects. These regressions are supplied in Table S12, where Model 5 corresponds to the Social Rank Composite model in Table S11. As above, descriptively, the intercept across all specifications is lower than .50 and skewed towards 0, indicating that subjects tend to choose the deepening voice recording as more descriptive of rank-related traits. In fact, the inclusion of additional controls *lowers* the predicted intercept, which goes from .41 in the baseline model with no predictors to .32 in the final model that includes all 4 predictors. In Models 1 to 4, the estimation of the intercept is rather precise, with standard errors ranging from .02 to .08; these intercepts differ significantly from .50 (*p*s = .0003, .0004, .018, .021. and .115). However, despite a comparable effect size, the intercept in Model 5 is not well estimated due to the substantially unbalanced nature of the recruitment method control (*n* = 93 Mechanical Turk Workers; *n* = 181 undergraduate students) and the smaller sample size of the Mechanical Turk subjects. The standard error of the intercept in this model jumps up to .12 and the intercept is significant only at the 11.53% level.

Taken at face value, these additional analyses suggest that subjects perceive the deepening voice as signaling greater rank seeking, compared to the rising voice. These checks suggest that it is unlikely that this effect is driven by methodological artifacts or subject characteristics.

TABLE S11. ORDINARY LEAST SQUARE REGRESSIONS PREDICTING CHOICE OF STIMULI (AVERAGED ACROSS RANK ITEMS, DOMINANCE ITEMS, AND PRESTIGE ITEMS) FROM PARTICIPANT-LEVEL CONTROL VARIABLES. COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

	Social Rank Composite	Dominance Composite	Prestige Composite
Gender	0.01	0.02	-0.05^{+}
(Male = -1, Female = +1)	(0.03)	(0.03)	(0.03)
Age	0.00	0.01^{+}	-0.00
	(0.00)	(0.01)	(0.00)
Order of Stimuli	0.04	0.06*	-0.05^{+}
(Lowering Pitch First = -1, Raising Pitch First = +1)	(0.02)	(0.03)	(0.03)
Recruitment Method	0.00	-0.03	-0.02
(Undergraduate Student = -1, Mechanical Turk Worker = +1)	(0.04)	(0.04)	(0.04)
(Intercept)	0.32^{a}	0.21^{b}	0.53^{c}
	(0.12)	(0.14)	(0.13)
AIC	270.00	367.63	320.03
BIC	288.06	385.70	338.09
N	274	274	274

$^{+}$ $p < 0.10$, * $p < 0.05$ (for H_0: $b = 0$)

Note. Wald tests were used to test whether the model intercept differs significantly from .50 (H_0: $b_0 = .50$), which indicates random choice. A model intercept closer to 0 indicates the tendency to select the deepening voice as more characteristic of the trait, relative to the rising voice. ^{a} $p = .115$, ^{b} $p = .035$, ^{c} $p = .795$.

TABLE S12. ORDINARY LEAST SQUARE REGRESSIONS PREDICTING CHOICE OF STIMULI (AVERAGED ACROSS RANK ITEMS). COEFFICIENTS ARE FOLLOWED BY STANDARD ERRORS IN PARENTHESES.

	Model 1	Model 2	Model 3	Model 4	Model 5
Gender		-0.00	0.01	0.01	0.01
(Male = -1, Female = +1)		(0.02)	(0.02)	(0.02)	(0.03)
Age			0.00	0.00	0.00
			(0.00)	(0.00)	(0.00)
Order of Stimuli				0.04	0.04
(Lowering Pitch First = -1, Raising Pitch First = +1)				(0.02)	(0.02)
Recruitment Method					0.00
(Undergraduate Student = -1, Mechanical Turk Worker = +1)					(0.04)

(Intercept)	0.41[a] (0.02)	0.41[b] (0.02)	0.31[c] (0.08)	0.31[d] (0.08)	0.32[e] (0.12)
AIC	266.54	268.54	268.66	268.00	270.00
BIC	270.16	275.77	279.50	282.45	288.06
N	274	274	274	274	274

Note. Wald tests were used to test whether the model intercept differs significantly from .50 (H_0: $b_0 = .50$), which indicates random choice. A model intercept closer to 0 indicates the tendency to select the deepening voice as more characteristic of rank seeking, relative to the rising voice. [a] $p = .0003$, [b] $p = .0004$, [c] $p = .018$, [d] $p = .021$, [e] $p = .115$.

SUPPLEMENTAL REFERENCES

Apicella, C. L., & Feinberg, D. R. (2009). Voice pitch alters mate-choice-relevant perception in hunter–gatherers. *Proceedings of the Royal Society of London B: Biological Sciences*, *276*(1659), 1077–1082. http://doi.org/10.1098/rspb.2008.1542

Bassiri, D. (1988). Large and small sample properties of maximum likelihood estimates for hierarchical linear models. *Unpublished Doctoral Dissertation*, Michigan State University, East Lansing.

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59.

Cartei, V., Cowles, H. W., & Reby, D. (2012). Spontaneous voice gender imitation abilities in adult speakers. *PLoS ONE*, *7*(2), e31353. http://doi.org/10.1371/journal.pone.0031353

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. http://doi.org/10.1037/a0030398

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition*, *117*(3), 319–331. http://doi.org/10.1016/j.cognition.2010.09.003

Ives, D. T., Smith, D. R. R., & Patterson, R. D. (2005). Discrimination of speaker size from syllable phrases. *The Journal of the Acoustical Society of America*, *118*(6), 3816–3822. http://doi.org/10.1121/1.2118427

Jones, B. C., Feinberg, D. R., DeBruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, *1*(79), 57–62. http://doi.org/10.1016/j.anbehav.2009.10.003

Kreft, I., & de Leeuw, J. de. (1998). *Introducing multilevel modeling*. Thousand Oaks, CA: Sage.

Mathieu, J. E., Aguinis, H., Culpepper, S. A., & Chen, G. (2012). Understanding and estimating the power to detect cross-level interaction effects in multilevel modeling. *Journal of Applied Psychology*, *97*(5), 951–966. http://doi.org/10.1037/a0028380

Nieuwenhuis, S., Forstmann, B. U., & Wagenmakers, E.-J. (2011). Erroneous analyses of interactions in neuroscience: A problem of significance. *Nature Neuroscience*, *14*(9), 1105–1107. http://doi.org/10.1038/nn.2886

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, *129*(4), 2201–2212. http://doi.org/10.1121/1.3552866

Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. C. (2007). Men's voices as dominance signals: Vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, *28*(5), 340–344.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–164. http://doi.org/10.2307/271063

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Newbury Park, CA: Sage.

Rogosa, D. R., & Willett, J. B. (1983). Demonstrating the reliability the difference score in the measurement of change. *Journal of Educational Measurement*, *20*(4), 335–343. http://doi.org/10.1111/j.1745-3984.1983.tb00211.x

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods*, *12*(2), 347–367. http://doi.org/10.1177/1094428107308906

Sinnott, J. M., Owren, M. J., & Petersen, M. R. (1987). Auditory frequency discrimination in primates: Species differences (Cercopithecus, Macaca, Homo). *Journal of Comparative Psychology*, *101*(2), 126–131. http://doi.org/10.1037/0735-7036.101.2.126

Smith, D. R. R., Patterson, R. D., Turner, R., Kawahara, H., & Irino, T. (2005). The processing and perception of size information in speech sounds. *The Journal of the Acoustical Society of America*, *117*(1), 305–318.

Snijders, T. A. B. (2005). Power and sample size in multilevel linear models. In B. S. Everitt & D. C. Howell (Eds.), *Encyclopedia of statistics in behavioral science* (Vol. 3, pp. 1570–1573). Chichester, England: Wiley.

Snijders, T. A. B., & Bosker, R. J. (1993). Standard errors and sample sizes for two-level research. *Journal of Educational Statistics*, *18*(3), 237–259. http://doi.org/10.2307/1165134

Snijders, T. A. B., & Bosker, R. J. (1994). Modeled variance in two-level models. *Sociological Methods & Research*, *22*(3), 342–363. http://doi.org/10.1177/0049124194022003004

Stel, M., Dijk, E., Smith, P. K., Dijk, W. W., & Djalal, F. M. (2012). Lowering the pitch of your voice makes you feel more powerful and think more abstractly. *Social Psychological and Personality Science*, *3*(4), 497–502.

van der Leeden, R., & Busing, F. M. (1994). First iteration versus final IGLS/RIGLS estimators in two-level models: A Monte Carlo study with ML3. *Leiden, the Netherlands: Leiden University*.