Scripps Senior Theses   Scripps Student Scholarship

2022

Pitch Perfect: Impression Formation and Impression Management in Women's Pitch Modulation

Grace A. Corrigan
Scripps College

Follow this and additional works at: https://scholarship.claremont.edu/scripps_theses

Part of the Psychology Commons

PITCH PERFECT: IMPRESSION FORMATION AND IMPRESSION MANAGEMENT IN WOMEN'S PITCH MODULATION

By
GRACE CORRIGAN

SUBMITTED TO SCRIPPS COLLEGE IN PARTIAL FULFILLMENT OF THE DEGREE OF BACHELOR OF ARTS

PROFESSOR MA
PROFESSOR FOUGHT

APRIL 22, 2022

Acknowledgements

First and foremost, I would like to extend my utmost gratitude to Professor Jennifer Ma, who met with me weekly and talked me through endless rewrites and edits to my thesis. She has been a rock for me at every stage since I first conceived of this topic more than 12 months ago in her Social Psychology class. Professor Ma is the finest educator I've ever had the privilege to learn from. It is because of her that I feel confident that this thesis is truly my best work.

A sincere thank you as well to Professor Carmen Fought, my second reader, for providing a much-needed sociolinguist's perspective as I tried to incorporate as much of my Linguistics minor as possible into my Psychology thesis. Professor Fought has been unwaveringly optimistic about my capabilities and never failed to boost my morale as needed.

To my parents, who listened to me vent about thesis for over a year: you kept me going. Mom, you talked me out of panic after panic when I didn't trust myself to meet each deadline or make this paper the absolute best it could be. Dad, you gave me strength every time you told me how proud you were. Thank you both from the bottom of my heart for not only putting me through college but helping me thrive here.

To my friends: thank you. You were simultaneously my motivation in crunch time, my comic relief and shoulder to cry on, and my most vocal supporters.

My sincere gratitude to you all. This thesis would not exist without each and every one of you.

Abstract

How does the pitch of a woman's voice impact how she is perceived, and how might women change the pitch of their voices to fit the situation at hand? Study 1 examined whether pitch plays a role in impression formation. Participants listened to two women's voices at three pitch levels (raised, unchanged, lowered) and rated the speakers' personality traits. Ratings of speaker competence, confidence, and intelligence were significantly lower for the pitch-raised voices than for the unchanged or pitch-lowered voices. Additionally, ratings of speaker persuasiveness and attractiveness were significantly lower for the pitch-raised voices than for the unchanged voices. No effect of pitch on sociability ratings was observed, but ratings of femininity were significantly lower for the pitch-lowered voices than for the unchanged or pitch-raised voices. Study 2 investigated whether women would modulate their pitch in different conversational contexts. Female participants were recorded answering questions in neutral, flirtatious, and professional conversational contexts over Zoom. No effects of context were observed for participants' minimum, maximum, and median pitch, but participants' mean pitch was significantly lower in the professional context than in the neutral context. The results of these studies suggested that pitch may be a factor in the formation of impressions about female speakers, and that women may, whether or not they are aware of the role of pitch in impression formation, modulate their voices to appear more professional.

Pitch perfect: Impression formation and impression management

in women's pitch modulation

In 2003, Elizabeth Holmes, a young woman with an unusually deep voice, founded a health technology company called Theranos. Its product could supposedly perform cheap, expedited blood tests with just a few drops of blood from a customer's finger. Theranos grew quickly, and in 2015 *Forbes* labelled Holmes the youngest female self-made billionaire in America (Herper, 2016). However, soon after Holmes had received the accolade, investigators found that Theranos had not published any peer-reviewed research and that most of its claims were exaggerated. By 2018, numerous lawsuits had been filed against Theranos and Holmes for allegedly falsifying information about the company's products (Griffith & Woo, 2022). As of January 3rd, 2022, Holmes has been found guilty of three counts of wire fraud and one count of conspiracy to commit wire fraud, with a possible sentence of up to 20 years in prison (Griffith & Woo, 2022).

How did Holmes successfully con so many people for so long? Some theorize that her Steve Jobs lookalike aesthetic—Holmes frequently donned Jobs' classic black turtlenecks—made her investors associate her with the deceased Apple founder (Hallemann, 2019). One of Holmes's most striking qualities, though, is her unusually deep voice. She speaks in a register far lower than the typical range for cisgender women, though it is not impossible; some women's voices naturally fall in Holmes's range. However, several people from Holmes's past have stepped forward to say that the voice she uses in her public appearances is much lower than her natural speaking voice (Hallemann, 2019). Although Holmes's family maintains that she has spoken this way for years, one of Holmes's former professors asserts that Holmes spoke in a considerably higher register when she was a college student, and a former Theranos colleague of

Holmes's claims that Holmes occasionally slipped out of her low voice during the workday (Hallemann, 2019).

It would be difficult to confirm whether or not Holmes's voice is naturally as deep as it sounds, but she may have recognized the potential benefits of speaking in a lower register. People with lower-pitched voices are often perceived as more competent (Guyer et al., 2018; Krahé et al., 2021; Oleszkiewicz et al., 2017; Sorokowski et al., 2019), persuasive (Guyer et al., 2018), and trustworthy (Apple et al., 1979; Oleszkiewicz et al., 2017; Streeter et al., 1977). As a young woman in a male-dominated field, Holmes needed to project an image of a strong leader whose claims could be taken seriously. With her outfits, she channeled Steve Jobs; with her voice, she channeled a smart, competent woman who had invented a miraculous blood-testing product. Holmes's carefully crafted professional persona, possibly including her speaking pitch, was apparently all a façade. Do other women change their voices to project a certain image? What differing impressions might emerge based on variations in women's pitch levels?

Impression Formation

Although Holmes primarily utilized her voice and style to influence how she was perceived by others, people may consider many different cues when they form an impression of someone. Impression formation happens quickly, but people's initial opinions tend to last (Black & Vance, 2021; Digirolamo & Hintzman, 1997). Before someone goes to a job interview, a new school, or another novel environment, they are usually warned that first impressions matter. They might go to great lengths to present themselves well in numerous ways, including through their clothes, their attitude, and even their voice. However, social psychological research indicates that first impressions are typically based on only a few salient details, rather than on all of the information available about a particular person. According to Taylor (1981), humans are

"cognitive misers" who utilize the minimum amount of information needed for an impression

formation process. The goal is to form an accurate impression with minimal effort, so to save

energy and time, human brains selectively incorporate certain characteristics into their

impression of another person. The field of social psychology has proposed several impression

formation mechanisms that align with Taylor's cognitive miser theory.

Anderson's (1962) information integration theory (IIT) suggests that impression

formation is a process that takes numerous factors into account at once and considers some

factors more prominently than others. These factors might include a person's behavior, physical

characteristics, or even their speaking voice. In the research presented here, pitch will be treated

as a factor that might be considered in IIT. For example, a particular person might give more

value to other people's pitch than to their height. IIT would say that a short man with a deep

voice might produce a more favorable first impression for the observer than if the man were tall

and had a higher voice. According to IIT, impressions are formed from the aggregation of

various pieces of information, and these pieces of information have different "weights." In the

example above, pitch would have a larger weight than height and thus be given more

consideration during the impression formation process, so that when all traits were integrated,

pitch would be more impactful in determining the overall impression.

A second potential mechanism of impression formation might involve priming, the

cognitive process whereby exposure to some stimulus activates a certain construct (Ansorge,

2010). Studies of priming and impression formation have often examined stereotyping and

assumptions. For example, Banaji et al. (1993) found that exposure to a prime relating to

dependence, a stereotypically female trait, caused participants to rate an ambiguous female target

as being more dependent than an ambiguous male target. The same effect did not occur in

participants exposed to a neutral prime. Unlike IIT, which considers traits independently of one another during impression formation (Anderson, 1962), priming activates a specific concept or idea that may in turn activate other related constructs (Ansorge, 2010).

Silverstein (2003) described a language-specific phenomenon similar to priming: indexical order. This concept has been traditionally applied to studies of dialectic variation and has been used to examine linguistic variables that "index" (i.e., signal) a person's identity in some linguistic community. Typically, use of some dialect-specific variation is referred to as a "first-order index," which simply serves to position the speaker as a member of that dialectic community (Silverstein). For example, hearing a speaker's accent may indicate to listeners that they are a year-round inhabitant of Martha's Vineyard (a common sociolinguistic community of study). The first-order index in turn activates the "second-order" index, which comprises subconscious beliefs and characteristics associated with the speaker's dialectic community (Silverstein). In the Martha's Vineyard example, the speaker's accent might also cause the listener to subconsciously associate the speaker with stereotypes about year-round islanders.

Although indexicality usually refers to dialectic variables, it can be applied to pitch. A woman's voice first informs the listener that she is female. The speaker's gender, as indexed by her pitch (which, with some exceptions, would presumably fall in the normal range for female speakers and would not be mistaken for a male voice), would be the first-order index. The pitch of the speaker's voice relative to other speakers of the same gender would then activate a series of other ideas related to gender. A low female speaking voice may index higher masculinity or lower femininity relative to other women. It may also, separately, index traits typically associated with masculinity, such as competence (Krahé et al., 2021; Oleszkiewicz et al., 2017; Sorokowski et al., 2019) and confidence (Guyer et al., 2018). Though these traits are related to

masculinity, they are not its equivalent. As such, hearing a low female voice may not directly

activate masculinity at all, but assumptions about the speaker's competence and confidence

could still arise. Essentially, a woman with a low voice might not be considered outright

masculine, but she might be assumed to be in possession of more supposedly masculine traits

than other women. Similarly, a high female voice might not overtly prime ideas about the

speaker's femininity—though it could—but it may activate assumptions about the speaker's

possession of traits associated with femininity. Each of these possible characteristics

(masculinity, competence, etc.) would be activated via second-order indexicality. Notably,

Eckert (2008), in discussing Silverstein's indexicality theory, asserts that indexicality is not

linear, despite what its numerical order would suggest. Instead, first-, second-, and nth-order

indexicalities occur simultaneously (Eckert), such that a speaker's gender and her supposed

qualities are indexed all together via her pitch. Indexicality serves as a form of priming that

specifically functions to situate a speaker in a specific linguistic community—or, in the case of

the present research, as the target of a set of assumptions formulated based on the speaker's

pitch.

Priming has been shown to have a significant impact on impression formation (Higgins et

al., 1977; Srull & Wyer, 1979). Taking note of a certain quality might activate ideas about other

potential personality traits. One of the first things people may notice when they meet a new

person is their voice, and by extension, their pitch. Pitch, or the voice's fundamental frequency

(F_0), refers to the rate of vibration of the vocal folds. Faster vibration produces a higher pitch at a

shorter wavelength, while a lower pitch emerges from slower vibration (Titze, 2000). Although

humans can produce a vast spectrum of fundamental frequencies, cisgender men typically have a

significantly lower pitch range (on average 125 Hertz (Hz)) than do cisgender women (on

average 200 Hz; Titze, 2000). This difference makes pitch a key factor in indexing a person's

gender, but pitch may also invite further assumptions via second-order indexicality.

Evolutionary Psychology and Pitch

One of the primary reasons that people might consider pitch during impression formation

could come from the role of pitch as an indicator of sexual fitness. Evolutionary psychology

posits that many of the regular psychological processes humans experience are derived from

sexual selection, the way animals and people select the best partners with whom to propagate

their genetic material (Carmen et al., 2013). Today, humans do not have the same need to

quickly find and mate with a reproductive partner as other species did and do. With modern

conveniences like online dating and the ability to travel longer distances with ease, people can

afford to be choosier and take more than just reproductive fitness into account. However, the

evolutionary mechanisms designed to facilitate the process of passing down one's genes still

exist in human brains. During impression formation, people may consciously or unconsciously

consider characteristics that they perceive to be related to evolutionary fitness. Qualities like

narrow waists and large breasts in women are positively correlated with reproductive potential

(Jasieńska et al., 2004) and are accordingly considered attractive secondary sexual characteristics

cross-culturally (Singh et al., 2010). The development of many such secondary sex

characteristics is linked to estrogen in women (Jasieńska et al., 2004).

The majority of secondary sex characteristics begin to develop during puberty (between

the ages of 8 and 14). Like the rest of the body, human vocal cords also undergo changes during

this time. In men, increased testosterone in puberty causes the vocal cords to thicken and widen,

producing a noticeably deeper voice (about an octave lower than a child's voice) that signals

sexual maturity (Abitbol et al., 1999; Hari Kumar et al., 2016). Men also develop an Adam's

apple, another secondary sex characteristic. Although women's vocal cords also thicken during

puberty, their voices do not change nearly as much as men's do, as their vocal cords are still

quite narrow compared to men's (Abitbol et al., 1999). Women's pitch typically decreases by

about three tones during puberty, again signaling sexual maturity. In adulthood, female voice

pitch varies slightly as estrogen and progesterone levels fluctuate, primarily in line with different

phases of the menstrual cycle (Abitbol et al., 1999; Banai, 2017). In a study of women's vocal

characteristics at the late follicular, luteal, and menstrual phases of their cycles, Banai (2017)

found that women's minimum pitch was significantly higher during the late follicular (most

fertile) phase than in the menstrual phase, suggesting that higher pitch might be an indicator of

fertility. Abitbol et al. (1999) also noted that participants' estrogen levels increased steadily

during the late follicular phase, indicating that pitch and estrogen may be positively related. By

contrast, in men, and in postmenopausal women, levels of androgens such as testosterone are

negatively related to pitch (Abitbol et al., 1999; Hari Kumar et al., 2016).

Most people are not aware of these relationships between vocal pitch and sex hormones;

however, evolutionary psychology would suggest that people may still associate pitch with

femininity and masculinity. A woman's high-pitched voice may signal, albeit unconsciously, that

she has higher estrogen levels than a woman with a lower voice. In turn, the listener may

experience priming of traditionally feminine characteristics, such as general femininity and

warmth. Likewise, a low male voice may signal that the speaker has higher testosterone levels

than a man with a higher voice, priming traditionally male characteristics like competence and

dominance. It may be the case that sex-atypical voices, such as low female voices and high male

voices, are less easily interpreted in impression formation. They may simply be less strongly

linked with the traits typically associated with the speaker's gender, or they may outright prime

traits associated with the opposite sex. As discussed earlier, indexicality theory would suggest that it is not even necessary for masculinity and femininity themselves to be overtly primed for other, associated traits to be indexed by pitch.

Pitch & Perceptions of Trustworthiness

The earliest research on pitch as a cue used during impression formation was conducted by Streeter et al. (1977), who investigated whether voice pitch might drive perceptions of negative traits, such as dishonesty. In their first experiment, male dyads acted as interviewer and interviewee, and interviewees were instructed to falsify half of their responses. The researchers found that the speakers' voice pitch increased by 3.3 Hz on average when their responses were false. In a second experiment, participants were presented either with unaltered audio from the honest and dishonest conditions from the first experiment, or with a filtered version that distorted semantic content but left the interviewees' pitch variation intact. The results indicated that participants' truthfulness ratings were significantly negatively correlated with the average pitch of interviewee responses in the filtered audio condition. Notably, no such correlation appeared in the unfiltered audio condition, suggesting that people might not typically treat pitch as an indicator of dishonesty unless it were the only cue available to them. Building on Streeter et al., Apple et al. (1979) studied the effects of speakers' voice pitch and speech rate on ratings of truthfulness and other personal attributions. Participants heard a series of pitch-raised, pitch-lowered, and unmanipulated male voice recordings. They rated the lowest-pitched recordings the most truthful, followed by the unmanipulated and finally the highest-pitched speech.

In later years, other researchers continued to examine the relationship between pitch and trustworthiness. O'Connor et al. (2011) looked at infidelity, an example of untrustworthy behavior in a mating context. Male and female participants listened to and rated male and female

voice stimuli. Speakers' voice pitch was modified, creating a raised/lowered pitch pair for each

speaker, and participants chose which of the two voices belonged to a speaker who was more

likely to cheat on their partner. Based on the proportion of times in which the lower voice was

selected, O'Connor et al. found that women believed lower male voices were more likely to

cheat than higher male voices. By contrast, male raters believed that higher female voices were

more prone to infidelity than were lower female voices.

Given that studies like Apple et al. (1979) and O'Connor et al. (2011) indicated that

people may assess speakers' trustworthiness based on their voice pitch, Schild et al. (2020)

sought to determine whether voice pitch was a reliable predictor of men's actual, in addition to

perceived, trustworthiness. Male speakers completed several measures of their own

trustworthiness. Notably, only one significant relationship was observed between pitch and any

of the actual trustworthiness measures. A negative correlation emerged between mean pitch and

self-reported infidelity, such that lower-pitched speakers reported more instances of infidelity

than did higher-pitched speakers. A group of participants then listened to the men's voices and

completed several ratings of the speakers' trustworthiness. In line with the researchers' findings

about male infidelity, participants' judgments of mating-related trustworthiness were positively

correlated with mean pitch. By contrast, participants' perceptions of the speakers' economic

trustworthiness were negatively correlated with mean pitch, even though no such relationship

was observed between men's pitch and their actual trustworthiness. These findings suggest that

although participants seemed to treat it as such, pitch is not necessarily a valid indicator of men's

actual trustworthiness.

Oleszkiewicz et al. (2017) assessed blind and sighted adults' perceptions of

trustworthiness based on pitch. Irrespective of participants' sight ability, both male and female

voices received higher trustworthiness ratings when their pitch had been lowered. That blind and

sighted participants generally did not differ significantly in their ratings is an indication that both

groups process pitch similarly as a cue for making personality judgments. Indeed, these results

suggest that even sighted individuals might take pitch into account during impression formation,

despite having access to a wealth of visual information as well. The literature on pitch and

trustworthiness has analyzed this relationship from a variety of different perspectives and in

many different scenarios. In nearly every situation, however, lower-pitched male and female

voices are consistently perceived as being more trustworthy.

Pitch & Perceptions of Dominance

Another well-documented topic in the literature on pitch relates to perceptions of

dominance. Jones et al. (2010) found that for both male and female voices, pitch-lowered

versions were chosen as the more dominant significantly more often than pitch-raised versions.

In line with Jones et al.'s findings, Borkowska & Pawlowski (2011) found a negative linear

relationship between women's voice pitch and the dominance ratings they received from men

and women. In general, lowered female voices were perceived as more dominant than their

habitually pitched versions, while raised female voices were perceived as less dominant than

their habitually pitched versions.

Though most pitch research is based on others' perceptions of the speaker, some

researchers (e.g., Schild et al., 2020) have questioned whether a person's natural speaking pitch

relates to their own self-reported personality traits, including dominance. Using data from 11

previous studies, Stern et al. (2021) found that participants' self-reported dominance correlated

negatively with their mean pitch. These findings suggest that although the relationship is not

causal, feelings of dominance may be related to voice pitch. Because men and women with lower

voices tend to be perceived as more dominant (Borkowska & Pawlowski, 2011; Jones et al., 2010), it is possible that lower-pitched speakers are treated as though they are dominant, thereby influencing them to feel more dominant themselves.

Stern et al.'s (2021) research investigated people's natural speaking pitch, while studies like Borkowska & Pawlowski (2011) and Jones et al. (2010) electronically manipulated speakers' pitch. By contrast, Fraccaro et al. (2013) examined whether deliberately modifying one's voice would lead to differences in perceptions of dominance. Men and women recorded a series of vowels at their habitual pitch and at deliberately raised and lowered pitches, creating raised/habitual and lowered/habitual voice pairs. Female participants who listened to each voice pair chose the lower of the two voices as more dominant in all pitch pair and gender conditions. Male participants categorized the lower voices as more dominant in both male pitch pair conditions and in the female raised/habitual pairs, but not in the female lowered/habitual voice pairs. With the exception of men's ratings of female lowered/habitual pairs, Fraccaro et al.'s findings align with the rest of the literature on pitch and dominance. In general, speakers with lower voice pitch are typically perceived to be more dominant, a conclusion supported by Abitbol et al.'s (1999) findings about the relationship between testosterone and pitch in males.

Pitch & Perceptions of Attractiveness

Pitch research has been largely dedicated to studying the perceived attractiveness of various pitch levels, particularly for women's voices but also for men's. For instance, Fraccaro et al.'s (2013) study of intentionally modulated pitch and dominance also asked participants to select the more attractive of the two voices in each voice pair. All participants perceived men's habitually pitched voices as more attractive than their pitch-raised versions but showed no perceived differences between men's habitual and lowered voices. Among female participants,

women's habitually pitched voices were perceived as more attractive than their lowered and raised versions. However, among male participants, women's habitually pitched voices were only perceived as more attractive than their lowered versions. Like Fraccaro et al., Jones et al. (2010) examined attractiveness in addition to dominance. Both male and female participants found men's pitch-lowered voices more attractive than their pitch-raised versions. By contrast, only male participants showed a preference for pitch-raised female voices. Jones et al.'s results, and those of Fraccaro et al. (2013), are in line with those of Abitbol et al. (1999) about the relationship between pitch and sex hormones.

In another study, Zheng et al. (2020) divided male and female speakers into three groups based on their starting pitch: high, average, and low. Every speaker was then recorded speaking at their habitual pitch and at purposefully raised and lowered pitches, and participants selected the most attractive of each set of three pitch levels from the same speaker. For female voices at all three starting pitch levels, the pitch-raised version was judged more attractive than the other two pitch levels, and the habitually pitched version was judged more attractive than the pitch-lowered version. These findings are consistent with the literature on pitch and attractiveness. However, unlike the results seen for female voices, Zheng et al.'s results for male voices did not align with the patterns found by Fraccaro et al. (2013) and Jones et al. (2010). Male pitch-raised voices were considered more attractive than their habitual or pitch-lowered versions in Zheng et al.'s study. In fact, the pitch-lowered recordings from the average and low habitually pitched groups were rated the least attractive, and the pitch-raised versions from the low habitually pitched group were rated the most attractive. These findings are strikingly inconsistent with those of Fraccaro et al. (2013) and Jones et al. (2010), who found that lower male voices were chosen as the more attractive of the pair significantly more often than chance. It may be the case that

cultural perceptions of vocal attractiveness are different, because Zheng et al.'s research was conducted in China, whereas the other studies discussed here were conducted in Western countries.

Some of the research on vocal attractiveness has focused specifically on whether averageness is an appealing quality or not. A study conducted by Feinberg et al. (2008) investigated whether women's voices would be deemed more attractive if the pitch were approximately average for a female voice, or if attractiveness ratings increased as pitch increased. Men rated the attractiveness of a series of recordings from women whose voices ranged widely in pitch. The results showed a highly significant positive linear correlation between attractiveness ratings and pitch, so that as pitch increased, so did ratings of attractiveness. In a second experiment, female voices of low, average, and high starting pitch were raised and lowered, and the voices' attractiveness was rated by male and female participants. Men showed a stronger preference for high-pitched voices than women did. In fact, men selected the pitch-raised version of the voices in all three original pitch levels, whereas women only selected the pitch-raised version of low and average original voices, not high original voices. Overall, pitch raising had the strongest effect in increasing attractiveness ratings for low voices, and it had the weakest effect for high voices. Feinberg et al. concluded that, contrary to their predictions, people—particularly men—found high voices more attractive than average voices. These findings are somewhat inconsistent with Fraccaro et al. (2013), whose study indicated that habitually pitched female voices were more attractive than pitch-raised female voices.

Contrary to the findings of Feinberg et al.'s (2008) first study, Borkowska & Pawlowski (2011) did not observe a positive linear correlation between female voice pitch and male

participants' attractiveness ratings. Although low and medium starting voices that had been

pitch-raised were rated significantly more attractive than their unmodified and lowered versions,

high starting voices that had been pitch-raised were rated as less attractive than their unmodified

versions. Furthermore, a very high starting voice that underwent pitch-lowering was considered

significantly more attractive than its unmodified version. The pattern that emerged from

Borkowska & Pawlowski's research indicated a preference for medium-to-high female voices,

with a pitch threshold at which higher voices stop being more attractive. It is possible that

although a woman's higher voice may signal her fertility (Abitbol et al., 1999), too high a pitch

could come off as childish rather than sexually mature. The literature on pitch and attractiveness

has been largely inconclusive: does attractiveness increase as pitch increases, or are higher

voices sometimes less attractive than moderately pitched voices? Further research will be

necessary to establish how people perceive the attractiveness of female pitch.

Pitch & Perceptions of Competence

A substantial proportion of the literature dedicated to voice pitch has examined the

effects of pitch on perceptions of competence. For example, Oleszkiewicz et al.'s (2019)

investigation of blind and sighted individuals' ratings of vocal trustworthiness also showed that

pitch-lowered female voices were given higher competence ratings than pitch-raised female

voices. Another study, by Krahé et al. (2021), had participants listen to a pitch-raised and pitch-

lowered female voice, then rate the speaker on positive and negative feminine and masculine

attributes, competence, and likeability. The pitch-raised female voice received higher ratings on

both positive and negative feminine attributes than the pitch-lowered voice, although no

differences in masculinity ratings emerged based on pitch. Furthermore, the high-pitched female

voice received significantly higher likeability and lower competence ratings than the low-pitched

voice. The researchers then conducted the same experiment with a pitch-raised and pitch-lowered male voice. The low-pitched male speaker was rated as more masculine and less feminine, in terms of both positive and negative traits, than the high-pitched speaker. The low-pitched male speaker also received lower likeability ratings than the high-pitched speaker. However, the low-pitched male speaker was not rated significantly more competent than the high-pitched speaker. Krahé et al.'s findings were thus somewhat inconsistent between male and female speakers, because the female speaker's competence was rated lower when pitch was higher, while the male speaker's competence ratings did not appear to differ based on pitch.

In another study of competence, Sorokowski et al. (2019) investigated whether professional contexts led men and women to modulate their voices and whether speech samples from these professional contexts were perceived as being more authoritative and competent. In Sorokowski et al.'s first experiment, a group of academic professionals were asked to give a talk on their field of expertise (authority speech) and to give directions to a nearby building (control speech). The researchers found that authority speech had a lower mean pitch than control speech, and that the difference in speaker pitch between the neutral and authority conditions was significantly larger for women than for men. In Sorokowski et al.'s second experiment, non-Polish-speaking participants listened to recordings from the authority and control conditions, which had been conducted in Polish. Free of the influence of semantic content, participants judged recordings from the authority speech condition to be more competent and more authoritative than those from the control speech condition. Notably, this effect was more pronounced for male speakers than female speakers.

Sorokowski et al. (2019) and Krahé et al. (2021) demonstrated that adults typically associate lower pitch with higher competence. Seeking to identify whether younger participants

produced similar pitch-based judgements, Cartei et al. (2021) examined children's assessments

of adults' competence. Nine-year-old children listened to pitch-raised, pitch-lowered, and

unmodified female voices before rating the speakers' competence for stereotypically male,

stereotypically female, and gender-neutral occupations. For female voices, pitch-raised voices

received the highest competence ratings for the stereotypically female jobs and the lowest

competence ratings for the stereotypically male jobs. Notably, in no job condition did women's

pitch-lowered voices receive higher competence ratings than averagely pitched voices. Cartei et

al.'s findings are somewhat inconsistent with those of Krahé et al. (2021) and Sorokowski et al.

(2019), who found that lower female voices were generally rated more competent than higher

female voices. It may be that higher female voices were considered the most competent in

stereotypically female roles because women, and by extension more feminine-sounding voices,

are believed to be the best fit for these roles.

Another study, by Guyer et al. (2018), examined competence, persuasiveness, and

confidence. A male speaker was rated more confident and was shown to be more persuasive at a

lowered pitch than at a raised pitch. Ratings of intelligence were also shown to correlate strongly

and positively with lowered pitch, and in line with previous research (Krahé et al., 2021;

Oleszkiewicz et al., 2019; Sorokowski et al., 2019), Guyer et al. found that competence was

positively correlated with lowered pitch as well. In general, the literature on pitch and

competence has been fairly consistent, but little to no research has been conducted to investigate

the relationship between female pitch and ratings of confidence, persuasiveness, and intelligence.

Impression Management

The literature described thus far has looked at the role of pitch in impression formation,

an important social psychological process. Just one simple quality of a person's voice may cause

the speaker to be perceived in a variety of different ways or be ascribed several different

personality traits by listeners. Although there are downsides to this phenomenon, knowledge of

others' impression formation based on pitch may also have its benefits. Impression management,

in which a person modifies their behavior to create and maintain others' perceptions and

opinions of them (Goffman, 1959), can occur via voluntary pitch modulation. As was alluded to

earlier, Elizabeth Holmes may have purposefully lowered her voice to seem more competent and

trustworthy. She may have known, or believed, that women with lower voices are perceived as

such, and changed her voice pitch to align with the image she wanted to project.

Goffman (1959) considered impression management a theatrical process. In Goffman's

view, people are actors who perform for audiences in various environments. Both parties (actor

and audience) take cues from each other and from the environment to construct a definition of

their shared situation. Environmental cues include the setting and the context of the situation.

Cues from the audience may be favorable or disapproving, depending on the contextual

appropriateness of the actor's behavior. Based on the definition the actor has conceived thus far,

the actor modifies their behavior in the hopes of forming a positive impression. Once they have

received a response from the audience, the actor determines whether or not to continue their

behavior or try a new tactic. Impression management as described by Goffman may not happen

in scenarios where the actor does not closely consider their own actions. However, situations that

require higher self-monitoring are generally also situations in which impression management

occurs. These situations may be low-pressure, like meeting someone new, or more significant,

like a job interview. In scenarios that elicit impression management behaviors, people may

consider modifying the pitch of their voices as a technique for changing the audience's view of

them.

Pitch Modulation

Although there has not been much research conducted on people's pitch modulation as a form of impression management, the literature on this topic falls into two categories: pitch adjustment in a mating context and pitch adjustment in a professional context. As discussed earlier, Sorokowski et al. (2019) found that people's mean pitch was lower when they spoke about their area of expertise than when they gave directions to a nearby building. Given the findings of numerous studies (Guyer et al., 2018; Krahé et al., 2021; Sorokowski et al., 2019) that suggested lower-pitched voices are perceived as more competent, people's choice to lower their voices in professional scenarios may be sensible.

One of the only other studies of pitch modulation and professionalism, conducted by Leongómez et al. (2017), investigated whether participants would adjust their pitch in response to perceived differences in status. Participants were shown a prototypical face and an employee testimonial about a dominant boss, a prestigious boss, and a neutral boss, respectively. They verbally answered three hypothetical interview questions while viewing each of the three faces, before completing self-report scales for their own dominance and prestige. The results indicated that more dominant participants lowered their mean pitch more (relative to the neutral target) for the dominant and prestigious targets than did less dominant participants. More dominant participants also showed a trend towards exhibiting less pitch variation. All participants spoke with a relatively higher mean pitch towards the two high-status targets, particularly if the participants perceived themselves as lower in dominance. These pitch modulation effects were most pronounced in the two interview questions that asked about participants' personal attributes and interpersonal skills, but not in the interview question that asked for introductory information.

Studies of pitch modulation and professionalism, like Leongómez et al. and Sorokowski et al. (2019), have found that participants tend to lower their voices in professional contexts. Other pitch modulation research has focused on whether participants adjust their pitch according to the attractiveness of their conversation partner. For instance, Hughes et al. (2010) asked male and female participants to leave voicemails for three members of the opposite sex: one attractive, one unattractive, and one neutral. They were shown pictures of the voicemail recipients before they spoke, so that they would have an idea in mind of the person's appearance. Both men and women spoke at a significantly lower mean and median voice pitch to the attractive targets than to the unattractive targets.

Using the same voicemail methodology, Fraccaro et al. (2011) investigated whether women would modulate their voice pitch in response to seeing an attractive male face. Female participants were asked to leave a voicemail for a masculinized prototype face and a feminized prototype face. The participants' preference for feminine versus masculine faces was assessed in separate attractiveness-rating blocks. In line with Fraccaro et al.'s predictions, the difference in the participants' voice pitch between the two voicemail conditions was positively correlated with the strength of their preference for one type of face or the other. In general, the women spoke in a higher-pitched voice to the face type that they had indicated that they found more attractive. These results are strikingly inconsistent with those of Hughes et al. (2010) and indicate the literature's general lack of consensus about pitch modulation in flirtatious contexts.

Although the hypothetical-voicemail procedure is one of the more common ways to examine pitch modulation, it is also possible to investigate pitch changes in the real world. Pisanski et al. (2018) studied men's and women's pitch modulation in a speed-dating context, wherein male-female pairs spoke for six minutes before switching conversation partners. At the

end of each "date," each person indicated their preference for their partner by marking "yes" or

"no" in their personal speed-dating booklets. The researchers found that women spoke with a

higher mean pitch, higher maximum pitch, and higher pitch variability when talking to men they

preferred than when talking to men they marked as "no." Specifically, women increased their

mean pitch and their pitch variability significantly with men whom they preferred and who had

low desirability scores (the number of yeses the man received over the total number of dates he

had at the event). Notably, by contrast, women marginally decreased their mean pitch and pitch

variability with men they preferred who had high desirability scores. Pisanski et al. also

examined whether any aspect of pitch might predict desirability. Indeed, women with lower

minimum pitch were significantly more desired by men, a result that counters the findings of

Feinberg et al. (2008), Fraccaro et al. (2013), Jones et al. (2010), and Zheng et al. (2020). This

inconsistency in research on pitch and attractiveness may stem from the fact that participants in

the aforementioned studies only heard women's voices in isolation, rather than in a live speed-

dating context. The findings of Pisanski et al. (2018) further serve to complicate the literature's

lack of consensus about women's pitch modulation in flirtatious contexts.

The Present Research

Although the literature on voice pitch is relatively broad and varied, there remain a few

significant gaps. One of these is the dearth of research on impression formation that looks

specifically at female voices. Particularly in the realms of competence and persuasiveness, the

bulk of research on pitch has been limited to male voices. In fact, almost no research has been

dedicated to the persuasiveness of female voices, and even male voice pitch has hardly been

examined in terms of persuasiveness (Guyer et al., 2018). Studies of competence usually look at

male and female voices in tandem (Cartei et al., 2021; Krahé et al., 2021; Oleszkiewicz et al.,

2019; Sorokowski et al., 2019) or only at male voices (Guyer et al., 2018). Female voices, by contrast, are typically studied in the context of their attractiveness. Far more research (Borkowska & Pawlowski, 2011; Feinberg et al., 2008; Fraccaro et al., 2013) has been dedicated to the attractiveness of female voices than male voices, yet studies of attractiveness and women's pitch have produced conflicting results. For a trait so heavily studied in the literature on female voice pitch, the effect of pitch on perceptions of attractiveness has been surprisingly difficult to pinpoint.

Due in part to the field's focus on vocal attractiveness, the literature on female voice pitch has thus far been of a fairly narrow scope and has failed to consider perceptions of traits like sociability, confidence, intelligence, and persuasiveness. In fact, no prior study has examined sociability in relation to pitch for either male or female voices, nor has any prior study examined perceptions of intelligence, persuasiveness, and confidence for female voices. Additionally, only a few studies have examined the relationship between female voice pitch and perceptions of competence. Given that research using male voices has discovered significant relationships between pitch and perceptions of the aforementioned qualities, it is unwise to ignore or understudy women's voice pitch. Female voices may follow the same patterns as male voices or they may produce entirely different results, but if the field does not spend more time researching female voices, those patterns will remain undiscovered.

In light of these gaps in the literature, Study 1 investigated people's impression formation based on women's speaking pitch. Participants listened to two female voices at three pitches (lowered, unchanged, and raised) and rated each of the voices on a number of qualities (competence, confidence, persuasiveness, intelligence, attractiveness, femininity, and sociability). Although attractiveness has arguably been overstudied, the details of its relationship

to pitch remain inconclusive. Therefore, Study 1 addressed perceptions of attractiveness in addition to the other dependent variables. Participants were expected to rate pitch-lowered voices higher than unchanged voices, and unchanged voices higher than pitch-raised voices, on the dimensions of competence, confidence, persuasiveness, and intelligence. Pitch-raised voices were expected to receive higher ratings of attractiveness, femininity, and sociability than unchanged voices, and unchanged voices were expected to receive higher ratings on these three traits than pitch-lowered voices.

Though much research has been devoted to pitch and impression formation, the field of research on pitch modulation as a form of impression management is much smaller and more limited. Although several studies have examined pitch modulation in response to the attractiveness of conversation partners (Pisanski et al., 2018) or voicemail recipients (Fraccaro et al., 2011; Hughes et al., 2010), hardly any studies have looked at pitch modulation in formal contexts. Leongomez et al. (2017) and Sorokowski et al. (2019) are some of the only studies to have examined pitch modulation in response to a professional situation. Although both found some level of pitch lowering, their methodologies and general research questions were quite different from each other, and there have been so few studies conducted in this area that it would be difficult to say that their results are widely applicable.

Some inconsistencies have also emerged in the results of pitch modulation and impression management studies. For example, although both Hughes et al. (2010) and Fraccaro et al. (2011) utilized a hypothetical voicemail procedure, female participants in the former study lowered their voice pitch towards men they found more attractive, while women in the latter study raised their voice pitch towards attractive voicemail recipients. Pisanski et al.'s (2018) findings, in which women raised their voice pitch towards attractive conversation partners, more

closely align with Fraccaro et al.'s results. However, as with pitch modulation in professional

contexts, there has been so little research in the area of pitch modulation and attraction that the

results of these three studies do not conclusively show whether and how women change their

voices around attractive people. More research in this area will be needed in order to solidify the

field's understanding of pitch modulation in a mating context.

No study has thus far directly compared how people modify their pitch in professional

versus mating contexts. Study 2 closed this gap by putting participants in three different

conversational contexts (neutral, flirtatious, and professional), then comparing their pitch across

all contexts. Participants' voices were recorded and analyzed for several vocal characteristics.

They were expected to speak with lower pitch in the professional context than in the neutral

context, and in turn speak with lower pitch in the neutral context than in the flirtatious context.

STUDY 1

An experimental study examined whether the pitch of female speakers' voices influenced

participants' beliefs about the speakers' personality traits. Participants were exposed to the same

six female voices: two at a raised pitch, two at a lowered pitch, and two whose pitch was

unchanged from their normal registers. Participants were expected to rate pitch-lowered voices

higher than unchanged voices, and unchanged voices higher than pitch-raised voices, on the

dimensions of speaker competence, confidence, persuasiveness, and intelligence. The opposite

pattern was expected to emerge for ratings of speaker attractiveness, femininity, and sociability.

Method

Participants

To determine an appropriate sample size, a power analysis was conducted using

G*Power (Faul et al., 2007) assuming the large effect sizes found in previous studies (Cartei et

al. 2021; Fraccaro et al. 2013; Guyer et al. 2013; Krahé et al. 2021) on similar topics. For a desired power of 0.80, $\alpha = 0.05$, and a repeated-measures design with a three-level independent variable, the minimum sample size required was 7 participants. However, to account for the possibility of the study having only a small effect size, the goal sample size was 90 participants.

A total of 92 participants were recruited via Amazon Mechanical Turk (MTurk), a crowdsourcing website that allows researchers to recruit participants remotely. The sample was approximately 38 (41%) women, 48 (52%) men, and 6 (7%) non-binary individuals. Racially, the sample was comprised of 60 (65%) White, 10 (11%) South Asian, 7 (8%) African-American, 6 (7%) East Asian, 5 (5%) Native American/Alaska Native, and 4 (4%) Hispanic/Latino participants. They ranged in age from 20 to 77 years old, with a mean age of 35.3 years ($SD = 10.36$). Participants received $5 as compensation because the study took less than 15 minutes to complete.

Materials

Voice recordings. Five cisgender women with similarly pitched speaking voices were recruited. Each woman recorded herself reading a short sentence: "Hello, it's a pleasure to meet you." A pilot test with 10 participants was conducted to select the two voices that differed the least from each other. Participants rated each speaker's likeability, competence, confidence, and femininity on a 5-point Likert scale (1 – *none at all,* 5 – *a great deal*). They also stated whether they felt any of the five voices belonged to the same person, and they rated the similarity of each voice pair on a 5-point Likert scale (1 – *not at all similar,* 5 – *extremely similar*). The two women whose pair of voices received the highest similarity rating ($M = 3.40$, $SD = 1.35$) and whose individual voices received the most similar likeability ratings ($M_1 = 3.90$, $SD_1 = 0.99$, $M_2 = 3.70$, $SD_2 = 1.49$) were selected to be the voices used in Study 1.

To create the stimuli for Study 1, the two women recorded themselves reading a brief passage: "Hello, it's a pleasure to meet you. I work just down the street. I'm quite good at my job. Have a nice day." Each speaker supplied 10 recordings of the same passage. They were asked to minimize the amount of emotion and pitch variation present in their voices during the reading so that participants' ratings of speaker qualities would be purely based on the women's habitual pitch, not other qualities. Rather than using only one recording and changing its pitch to three levels, the three most similar-sounding recordings from each woman were selected to minimize the risk of participants discovering the manipulation.

The six audio recordings were analyzed and modified with Praat (Boersma & Weenink, 2019), a phonetic analysis program. One recording from each speaker had its median pitch lowered by 20% so that the two pitch-lowered audio stimuli had a median pitch of 167.61 Hz and 172.72 Hz. One recording from each speaker had its median pitch raised by 20% so that the two pitch-raised audio stimuli had a median pitch of 262.26 Hz and 268.71 Hz. The two clips selected to remain "unchanged" had noticeably different starting pitches (median pitch 210.09 Hz versus 226.35 Hz). Rather than leave them unchanged and risk participants rating them differently based on the clips' 16-Hz difference, each clip was modified slightly so that they both had a median pitch of 218 Hz. Lastly, the intensity (volume) of each clip was set to 70 dB so that volume would not be a factor in participants' ratings.

Perceptions survey. A survey was designed and administered anonymously via Qualtrics, a survey design website. The first page contained one of the six auditory stimuli, which were presented in a random order. After participants had listened to the stimulus, they were presented with a series of statements to assess their perceptions of the speaker's competence, confidence, persuasiveness, intelligence, attractiveness, femininity, and sociability

(e.g., "The speaker is intelligent."). Using a 5-point Likert scale (1 – *strongly disagree*, 5 – *strongly agree*), participants rated their level of agreement with each of the statements. Items had strong face and construct validity, given that the statements referred to the exact constructs being measured. The measure was also expected to have high convergent validity with the measures used by other pitch researchers, such as Krahé et al. (2021), because Krahé et al. used the same Likert scale statement-rating methodology. However, convergent validity was only a supposition and not based on correlational analyses.

Procedure

Study 1 was conducted entirely online. Participants were recruited via MTurk, and no restrictions were set on which MTurk workers could participate, provided that they were 18 or older. At the start of the Qualtrics survey, participants completed an online informed consent procedure. After consenting, participants were asked to transcribe a sample audio message that said, "I am participating in this study." This task was designed to ensure that the participants could hear audio from the survey well enough to continue participating. The participants were then presented with the six auditory stimuli in a random order. After listening to each audio clip, the participants provided their perceptions of the speaker's traits. The order of the statements was also randomized for every speaker. After rating all six speakers, participants were debriefed and compensated.

Ethical Considerations

Study 1 was minimal risk. Participants listened to brief, neutral audio clips and rated the speakers on various traits using a series of Likert scales. Given that these tasks were not more dangerous or upsetting than what the participants would encounter in their day-to-day lives, the procedure was considered minimal risk. However, some deception was necessary. Participants

were told that they were hearing the voices of six different women, when in reality they only heard two women's voices whose pitch had been set to three different levels. It was necessary that the participants think that they were hearing different people, so that they would not discover the purpose of the study and consciously base their ratings on the differences in speaker pitch. If they had become aware of the manipulation, their ratings might have changed. They were informed of this deception, and the reasoning behind it, during the debriefing process. Knowledge of the deception was not expected to cause distress or generate ill will, because the deception was minor and did not affect confidentiality or other risk factors.

In addition to the procedure being minimal risk, no protected or vulnerable populations were sampled from in this study. Participants also were not asked to provide any sensitive information. The personal information they were asked for was purely demographic: age, gender, and race/ethnicity. Furthermore, because Study 1 was conducted anonymously online via MTurk and Qualtrics, participants' data were completely anonymous. Their IP addresses were not collected or stored, and they were not asked to give their names or other identifying information besides demographics. However, if a participant chose to email the investigator using MTurk's email function, their real name and email became visible, but these pieces of information were not linked to their data in any way. All participant data collected were stored securely within the investigator's password-protected Qualtrics account on a password-protected computer, and the data were only accessible to the investigator.

Participation in Study 1 was intended to be truly voluntary. Participants were compensated at a fair but not coercive rate, and they were informed during the consent process that they were free to withdraw or discontinue participation at any time without repercussions. Although the participants did not experience any direct non-monetary benefit, this study should

have scholarly merit and potentially benefit society at large. For instance, Study 1 is one of the first studies to examine the relationship between women's voice pitch and perceptions of intelligence, sociability, persuasiveness, and confidence. Furthermore, the study attempted to confirm prior studies' results concerning the relationship between women's voice pitch and ratings of attractiveness, femininity, and competence. Knowledge of the effects of pitch on impression formation may serve women who wish to make a certain first impression. Because Study 1 was minimal risk, the potential benefits to the literature and to society were expected to outweigh the risks of participation.

Results

A composite variable was created for each dependent variable (femininity, competence, intelligence, persuasiveness, confidence, sociability, attractiveness). For example, ratings of speaker 1 and speaker 2's femininity at each pitch level were averaged to achieve three femininity scores for each participant: femininity ratings of the pitch-raised voices, femininity ratings of the unchanged voices, and femininity ratings of the pitch-lowered voices. Overall, the voices were given moderate ratings on the seven qualities, with mean ratings ranging from 2.98 to 4.30 across all seven speaker qualities[1] (see Table 1). A separate repeated-measures Analysis of Variance (ANOVA) was conducted to analyze the effects of pitch on each dependent variable. Bonferroni correction was used as a follow-up test as needed.

---

[1] The skewness and kurtosis of each dependent variable was examined. All variables fell within acceptable levels (± 2) of skewness and kurtosis, with the exception of ratings of speaker femininity. Ratings of speaker femininity at all three pitch levels were negatively skewed. However, because participants had been explicitly told that they would be listening to female voices, this ceiling effect was expected. It was decided for the sake of consistent data analysis that this negative skew would not be problematic, so non-parametric tests were not conducted.

Table 1.

Descriptive Statistics for Ratings of Speaker Qualities at Three Pitch Levels

Speaker Qualities	Pitch-lowered	Unchanged	Pitch-raised
Competence	3.91 (0.71)$_a$	3.75 (0.70)$_a$	3.32 (0.86)$_b$
Intelligence	3.96 (0.65)$_a$	3.81 (0.71)$_a$	3.39 (0.90)$_b$
Confidence	3.79 (0.84)$_a$	3.62 (0.89)$_a$	3.05 (1.07)$_b$
Persuasiveness	3.28 (0.85)$_{a,b}$	3.39 (0.84)$_b$	2.98 (1.09)$_a$
Attractiveness	3.30 (0.91)$_a$	3.65 (0.80)$_b$	3.57 (0.73)$_{a,b}$
Sociability	3.14 (0.99)$_a$	3.26 (0.87)$_a$	3.18 (1.00)$_a$
Femininity	3.96 (0.84)$_a$	4.21 (0.70)$_b$	4.30 (0.79)$_b$

Note. Mean (*M*) and standard deviation (*SD*) given for rating of speaker qualities across three pitches.

[a] Within a row, cells that do not share a subscript are significantly different from each other at *p* < .05.

Participants' ratings of speaker competence differed based on pitch, $F(2,182) = 23.46$, $MSe = 0.36$, $p < .001$, $\eta^2 = 0.21$. In partial support of the research hypothesis, ratings of competence were significantly lower for the pitch-raised voices than for the unchanged or pitch-lowered voices (see Table 1). However, contrary to the hypothesis, no significant difference in competence ratings emerged between the unchanged and pitch-lowered voices. As shown in Table 1, the pattern of effects found for intelligence ($F(2,182) = 20.66$, $MSe = 0.38$, $p < .001$, $\eta^2 = 0.19$) and confidence ($F(2,182) = 24.66$, $MSe = 0.56$, $p < .001$, $\eta^2 = 0.21$) were the same as those found for competence.

There was also a significant difference in participants' ratings of speaker persuasiveness based on pitch, $F(2,182) = 7.76$, $MSe = 0.54$, $p < .001$, $\eta^2 = 0.08$. In partial support of the

research hypothesis, ratings of speaker persuasiveness were significantly lower for the pitch-raised voices than for the unchanged voices (see Table 1). However, contrary to the hypothesis, ratings of persuasiveness for the pitch-lowered voices were not significantly different from ratings for the other two voices. Unexpectedly, the same pattern emerged for ratings of speaker attractiveness ($F(2,182) = 7.38$, $MSe = 0.41$, $p < .001$, $\eta^2 = 0.08$; see Table 1), even though ratings of attractiveness had been predicted to be significantly higher for the pitch-raised voices than for the unchanged voices, and significantly higher for the unchanged voices than for the pitch-lowered voices.

In opposition to the research hypothesis, there was no significant difference in participants' ratings of speaker sociability based on pitch, $F(2,182) = 0.62$, $MSe = 0.53$, $p = .540$, $\eta^2 = 0.01$. By contrast, participants' ratings of speaker femininity differed based on pitch, $F(2,182) = 8.20$, $MSe = 0.35$, $p < .001$, $\eta^2 = 0.08$. Contrary to the hypothesis, no significant difference in femininity ratings emerged between the pitch-raised and unchanged voices. However, in partial support of the research hypothesis, ratings of femininity were significantly lower for the pitch-lowered voices than for the unchanged or pitch-raised voices (see Table 1).

Discussion

The purpose of Study 1 was to investigate how the pitch of female speakers' voices impacted listeners' perceptions of the speakers. Specifically, participants were expected to rate pitch-raised voices higher than unchanged voices, and unchanged voices higher than pitch-lowered voices, on the dimensions of speaker attractiveness, femininity, and sociability. The opposite pattern was anticipated to emerge for ratings of speaker confidence, competence, persuasiveness, and intelligence. No hypothesis was fully supported by the data, but many of the predictions were partially supported. For instance, ratings of competence were higher for pitch-

lowered voices than for unchanged voices, in line with previous research that found that pitch-lowered female voices received higher competence ratings than pitch-raised female voices (Krahé et al., 2021; Oleszkiewicz et al., 2019). Besides partially replicating prior studies' results regarding pitch and competence, Study 1 was the first in the field to examine a relationship between female voice pitch and ratings of persuasiveness, intelligence, and confidence. Although some results emerged suggesting that women with higher-pitched voices were considered less persuasive, confident, intelligent, and competent, many predicted differences were not found to be significant.

It is possible that the research hypotheses were not fully supported because of the chosen pitch levels presented to participants. More extreme (though still realistic and natural-sounding) pitch-raised and pitch-lowered stimuli may have produced more striking results. Notably, though, the pitch manipulations conducted in Study 1 were in fact more extreme than those conducted in other studies (Borkowska & Pawlowski, 2011; Feinberg et al., 2008; Guyer et al., 2018), in which audio stimuli were raised and lowered by 20 Hz. By contrast, voices in Study 1 were raised and lowered by 20% of their median pitch, which came out to roughly 40-45 Hz in each direction. However, if the pitch of the auditory stimuli had varied even more significantly, there might have been a risk of participants figuring out the purpose of the study. It may also be the case that the same relationships found by Guyer et al. for male speakers may not hold for female speakers. Thus, a more thorough investigation of the relationship between female voice pitch and these constructs will be necessary in future studies.

In addition to being the first to study constructs like persuasiveness, confidence, and intelligence in relation to female voice pitch, Study 1 was the first to investigate the relationship between sociability and any gender's pitch. The hypothesis that higher-pitched speakers would

receive higher sociability ratings was based in the literature suggesting that higher-pitched female voices are considered more likeable (Krahé et al., 2021), but no significant findings emerged for ratings of sociability in Study 1. The constructs of likeability and sociability may be too dissimilar to be adequately compared. Krahé et al. found a relationship between pitch and likeability, but perhaps there is no such relationship between pitch and sociability to be discovered.

Although sociability did not produce significant findings, ratings of speaker attractiveness did show an interesting pattern: the unchanged-pitch speaker was rated more attractive than the pitch-raised speaker. It had been hypothesized, however, that attractiveness ratings would increase from the pitch-lowered voices to the unchanged voices and from the unchanged voices to the pitch-raised voices. Because the much of the literature on pitch and attractiveness has differed in whether high (Feinberg et al., 2008; Jones et al., 2010; Zheng et al., 2020) or moderate (Borkowska & Pawlowski, 2011; Fraccaro et al., 2013) female voices are perceived as more attractive, Study 1 was intended in part to resolve these inconsistencies. Ideally, the results would have fully supported either camp: studies that suggest that moderate pitch is more attractive than raised and lowered pitch, or studies that suggest that attractiveness increases as pitch increases. Study 1 was predicted to support the latter and came closer to supporting the former, but its results were far from the desired conclusive evidence for the attractiveness of either moderate or high pitch.

Notably, the results for ratings of femininity also did not perfectly align with the research hypothesis. The unchanged and pitch-raised voices were found to be significantly more feminine than the pitch-lowered voices, but the two higher pitch levels were not rated significantly differently from each other. Because the bulk of the research hypotheses were based in theory

about traditionally feminine attributes, it would have been ideal for each pitch level to have been perceived as significantly more or less feminine than the others. Instead, because not all differences in femininity were significant, it is possible that some of the other inconsistencies in Study 1's results stemmed from participants not associating the three pitch levels with more or less feminine characteristics. If participants did not find the pitch-raised voices more feminine than the unchanged voices, then they may not have been primed to consider traditionally feminine or masculine attributes when they heard each voice, nor would these attributes necessarily have been indexed by the speakers' pitch levels. As suggested earlier, one possible solution to this issue would be to increase the gap in pitch between each voice so that the differences between the voices would be more pronounced.

STUDY 2

An experimental study investigated whether women would modify their voice pitch as a method of impression management when exposed to three different conversational contexts (flirtatious, neutral, professional). Participants were expected to speak at a significantly higher pitch in the flirtatious context than in the neutral context, and at a significantly higher pitch in the neutral context than in the professional context.

Method

Participants

To determine an appropriate sample size for this study, a power analysis was conducted using G*Power (Faul et al., 2007) based on the large effect sizes found by Hughes et al. (2010) and Sorokowski et al. (2019) in studies on similar topics. For a desired power of 0.80, $\alpha = 0.05$, and a repeated-measures design with three levels of the independent variable (context), the

minimum sample size required was 12 participants. To increase the study power, however, the target sample size was 30-40 participants.

In total, 32 participants were recruited via convenience sampling at a liberal arts college consortium in Southern California. Online recruitment materials were posted on relevant collegiate Facebook groups, Discord servers, and email lists. Word of mouth was also employed as a recruitment method, and participants recruited in this manner were given the same information as participants recruited online. Participants were limited to cisgender women because this study was intended to examine the natural fundamental frequency of individuals who both identified as female and were born with female vocal cords. Also, to ensure that there was at least a possibility of participants feeling flirtatious in the flirtatious context as designed, participants needed to be sexually attracted to men. The sample was comprised of 10 (31.3%) East Asian, 10 (31.3%) White, 5 (15.6%) Hispanic, 4 (12.5%) mixed-race, and 3 (9.3%) South Asian participants. They were all college aged. Each received a $10 gift card as compensation.

Materials

Prescreening survey. Individuals who indicated an interest in participating completed a short Qualtrics survey. The survey first asked if the individual was 18 or older and if they were a cisgender woman. If either of these questions was answered with "no," the individual was informed that they were not eligible for participation. If the two questions were both answered affirmatively, the individual completed the rest of the survey. Amidst a series of filler questions (their school, their major and/or minor, the frequency with which they used Zoom, and whether they had their own Zoom account), they were asked if they experienced sexual attraction to men and if they experienced sexual attraction to women. If they indicated that they did not experience sexual attraction to men, they were contacted after they had completed the survey and were

informed that they were not eligible for participation. No restrictions on participation were set

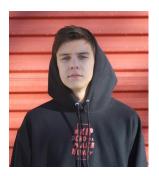based on individuals' sexual attraction to women.

Pre-recorded dialogue. Three college-aged men were hired to voice the three

conversation partners. Before data collection began, each man was recorded reading scripted

questions and responses associated with each conversational context so that there were

recordings of three different male voices reading all possible permutations of each context's

script.

General Zoom call characteristics. Participants were set up on a separate Zoom call for

each of the three contexts: neutral, professional, and flirtatious. On each call, the participant's

conversation partner's webcam was turned off, so that only his Zoom profile picture (see Figure

1) was visible. The neutral, professional, and flirtatious photos were of a young man who

appeared to be teenaged, a professional-looking man who appeared to be in his mid-twenties, and

a young man who appeared to be college aged, respectively. Photos were selected from a stock

photo website based on each man's approximate perceived age, as well as his facial expression

and positioning relative to the camera, both of which were intended to be fairly consistent across

the three men. The experimenter likewise ensured that the participant's webcam was off and then

left the room. Each conversation began shortly after the experimenter had left, as the

conversation partner's dialogue was simulated via the experimenter playing the recordings aloud

over Zoom. In each context, the participant's conversation partner started the conversation by

greeting the participant and introducing himself. He then asked the participant a total of three

questions designed to reflect the conversational context at hand. After the participant had

responded to a question, their conversation partner gave a brief scripted acknowledgement of

their response, then asked the next question. Once the participant had answered all three

questions, their conversation partner said goodbye and ended the Zoom call.

Figure 1.

Zoom Profile Pictures

Neutral context   Professional context   Flirtatious context

Note. Photographs from *Unsplash.com*, a website of freely-usable images (Buenrostro, 2020; del

Engelsen, 2018; Nate, 2019)

   Conversational context. There were three conversational contexts: neutral, flirtatious,

and professional.

   Neutral. Participants were told that they would be speaking with a high school junior,

"Matthew" (see Figure 1), who wanted to ask them about college. Matthew asked the following

three questions in order: "What is something about college that surprised you?", "What classes

are you taking this semester?", and "Besides graduating, what is your number one goal for your

remaining time at college?" Matthew's scripted speech was designed to mimic the speech of a

high school student. His script had a choice of responses (left to the discretion of the

experimenter) to the potential answers given by a participant. For example, depending on

whether the participant named a truly surprising thing about college or something run-of-the-

mill, Matthew either responded with "Oh, yeah, I never thought about that," or "That's what I've

heard; I'm kind of nervous."

Flirtatious. Participants were told that they would be speaking with a college student, "Quinn" (see Figure 1), who attended a nearby university. Quinn asked the following three questions in order: "If you could go anywhere right now, where would you go and why?", "Describe your perfect first date.", and "What are three things on your bucket list?" Quinn's scripted responses were designed to mimic the speech of a college student. For example, in response to the participant's answer to the second question, Quinn would say, "Hey, I could get into that. Sounds like a fun date." Quinn's questions afforded less opportunity for answers that required notably different responses, so Quinn only had one scripted response option corresponding to each of his questions.

Professional. Participants were told that they would be speaking with a job recruiter, "Adam" (see Figure 1), who would be asking them interview questions. Adam asked the following three questions in order: "What are some of your greatest skills, academically or otherwise?", "Where do you see yourself in five years?", and "What is a recent challenge you faced and how did you deal with it?" Like Matthew's script, Adam's script had a choice of responses (chosen from by the experimenter) to some of the potential answers given by participants. For example, depending on whether or not the participant had a five-year plan, Adam would either say, "It sounds like you've got a pretty good handle on what you want to do. That's great to hear," or "It's okay not to have everything figured out yet. College is a great time to start solidifying what you're passionate about."

Manipulation check. Participants were given a short survey to complete after they had experienced all three contexts. They were asked to rate how flirtatious and professional they found each of the three contexts, using a 5-point Likert scale (1 – *not at all*, 5 – *very much*). These scales were designed to see if the participants had noticed a change in context.

Inter-conversation questions. Participants were given a series of four questions to complete after each conversation had concluded. Using a 5-point Likert scale (1 – *not at all*, 5 – *very much*), they rated how clear the audio was on their most recent Zoom call, how comfortable they were with the questions that they had been asked, how comfortable they were with their most recent conversation partner, and how pleasant their most recent conversation was. These items served as fillers to give the experimenter time to switch their Zoom screen name and profile picture to those appropriate for the next context.

Positive reminiscence exercise. After they had been debriefed, participants completed a brief positive reminiscence exercise on paper. Though positive reminiscence exercises have existed for years in the positive psychological literature (Aşiret & Kapucu, 2016; Bryant et al., 2005), the exact prompt used in this study was designed by the experimenter. The prompt said, "Consider an event from your past that evokes joy. Remember as many details as you can, focusing on the pleasant feelings you felt in the moment. Take a few minutes to write down what you recall." Participants were given blank space below the prompt in which to respond, and there was no time limit.

Procedure

Upon arriving at the lab setting, each participant signed an informed consent document. They were told that the purpose of the study was to investigate how women interacted with men over Zoom. The experimenter also obtained verbal consent to record the participant and proceeded to do so. The participant was first set up on the neutral context Zoom call by the experimenter, who then left the room. The experimenter, now stationed in another room, began the conversation by playing a recording greeting the participant as "Matthew." After each of the three questions, the experimenter either selected and played the most appropriate scripted

response option or played the only scripted response option for that particular question and its

answer. When the neutral conversation had concluded, the participant was randomly assigned to

experience the flirtatious or professional context first before experiencing the other. The

question-answer procedure was repeated, using each context's respective recorded scripted

speech, for all three contexts. The order of the three men's voices was also randomized for each

participant so that each participant heard all three men, each reading a different role. In total,

there were 12 possible voice assignment and context order combinations. The participant

completed the inter-conversation questions after each call. After the participant had experienced

all contexts, they completed the manipulation check. The experimenter then debriefed the

participant, including an explanation about the experiment's deception, and asked them to

complete a short positive reminiscence exercise to counteract the effects of being deceived.

Lastly, the participant was compensated.

Ethical Considerations

Study 2 was minimal risk. Participants were recorded during three conversational

contexts over Zoom in a lab setting. The topics discussed in the three contexts were not any more

upsetting or personal than what might be discussed in the participants' day-to-day lives.

Although the study procedure was minimal risk, some deception was necessary. The

participants were deceived as to the identity of the men they spoke with on Zoom in each

context. In each scenario, the participants did not have conversations with their video cameras

on, so they only heard the voice of their conversational partner and saw his profile picture. For

the purposes of the study, photos of a boy who appeared to be high school aged, a man who

appeared to be college aged, and a man who appeared to be in his mid-twenties were used so that

the participants perceived the context as neutral, flirtatious, or professional, respectively. Three

men's voices were prerecorded reading scripts associated with each of the three conversational partners, and the recordings were played by the experimenter in each context. The men in the photos were not the same men who recorded the scripts. However, the participants needed to believe that they were speaking with the men in the photos in real time so that any impression management techniques they used, such as pitch modulation, would be due to the conversational context. They were informed of this deception and the reasoning behind it during the debriefing process. Knowledge of the deception was not anticipated to cause distress or generate ill will, because the participants knew that the experimenter would have heard their responses to the questions eventually (as part of the data analysis process). As an extra precaution, though, the debriefing process included a positive psychology exercise to improve participants' overall mood state. Prior research has suggested that positive reminiscence can decrease depression (Aşiret & Kapucu, 2016) and boost happiness levels (Bryant et al., 2005). This short exercise was assumed to be sufficient for reducing participants' possible mild distress in response to being deceived.

To further minimize potential risks, no protected or vulnerable populations were sampled from in this study, and participants were not asked to provide any sensitive information. However, their voices were recorded. Participants were warned of the recording process when providing informed consent so that they could withdraw ahead of time if they felt uncomfortable with being recorded. Because voices are considered identifiable information, and because the participants were present in a lab setting and were seen by the investigator, the data collected in Study 2 were confidential, not anonymous. However, the participants' names were not collected, and their voice recordings were heard only by the investigator. All participant data were stored securely on the investigator's password-protected computer, and the data were only accessible to the investigator.

Lastly, participation in Study 2 was intended to be truly voluntary. Participants were compensated at a fair but not coercive rate, and they were informed during the consent process that they were free to withdraw or discontinue participation at any time without repercussions. Although the participants did not experience any direct benefit, this study should have scholarly merit. For example, no prior study on pitch and impression management had directly compared women's pitch modulation in flirtatious contexts to their pitch modulation in professional contexts. In fact, hardly any studies had examined women's pitch modulation in professional contexts at all. Study 2 investigated this phenomenon. Additionally, previous research had produced inconclusive, conflicting results on how women adjust their voices in flirtatious contexts. Study 2 may add to this subsect of the literature by demonstrating that women raise the pitch of their voices in response to attractive conversation partners. Because Study 2 was minimal risk, the potential benefits to the literature were expected to outweigh the risks of participation.

Results

Manipulation Checks

The data from the manipulation checks were analyzed to determine whether there were significant differences in participants' ratings of how flirtatious and professional each conversation was. Separate repeated-measures ANOVAs were conducted to examine differences in flirtatious and professional ratings across the three contexts. Bonferroni correction was used as a follow-up test as needed.

There was a significant difference in participants' perceptions of the conversations' flirtatiousness, $F(2,62) = 115.04$, $MSe = 0.46$, $p < .001$, $\eta^2 = 0.79$. Specifically, in line with the intended research design, the flirtatious context ($M = 3.30$, $SD = 0.96$) was rated significantly

more flirtatious than the neutral ($M = 1.13$, $SD = 0.55$) and professional ($M = 1.03$, $SD = 0.18$) contexts, which were not significantly different from each other. Notably, though, although the flirtatious context was perceived as being more flirtatious than the other two, it was only rated somewhat flirtatious.

There was also a significant difference in participants' perceptions of the professional quality of the three conversations, $F(2,62 = 140.98$, $MSe = 0.56$, $p < .001$, $\eta^2 = 0.82$. In line with the intended research design, the professional context ($M = 4.78$, $SD = 0.49$) was rated significantly more professional than the neutral ($M = 2.94$, $SD = 0.98$) and flirtatious ($M = 1.66$, $SD = 0.79$) contexts. Unexpectedly, the neutral context was also rated significantly more professional than the flirtatious context.

Dependent Variables

Each participant's recording was analyzed using Praat (Boersma & Weenink, 2019). Their answers to each question were selected and set apart from the rest of the recording, then analyzed for their mean, median, minimum, and maximum pitch. On first analysis, there seemed to be considerable background noise, resulting in maximum pitch readings of around 490 Hz and minimum pitch readings of around 60 Hz. Therefore, to limit the inclusion of extraneous background noise, a restriction was set so that only pitches between 100 and 400 Hz were included in the analysis. A composite variable was created for each dependent variable in each context. For example, participants' median pitch across the three questions in the flirtatious context was averaged. Tests of normality were conducted on each composite variable. Because only a few composites showed slight normality issues[2], the data did not undergo any

---

[2] The skewness and kurtosis of each composite dependent variable was examined. Most variables fell within acceptable levels (± 2) of skewness and kurtosis, and those that did not had kurtosis ≤ 5.23 and skewness ≤ 3.47. Almost no Kolmogorov-Smirnov or Shapiro-Wilk tests of normality came back significant, suggesting that even these somewhat skewed and kurtotic distributions did not significantly violate assumptions of normality. The skewness and kurtosis values in this study were therefore not considered significant enough to necessitate transformation or use of a non-parametric test.

transformations. Separate repeated-measures analyses of variance (ANOVAs) were conducted to examine differences in participants' mean, median, minimum, and maximum pitch between the neutral, flirtatious, and professional contexts. Bonferroni correction was used as a follow-up test as needed.

There was a significant difference in participants' mean pitch across contexts, $F(2,60) = 4.25$, $MSe = 25.38$, $p = .019$, $\eta^2 = 0.12$. In support of the research hypothesis, participants' mean pitch was significantly lower in the professional context than in the neutral context (see Table 2). However, contrary to the research hypothesis, no other significant differences in mean pitch emerged between contexts.

Table 2.

Participants' Pitch Characteristics Across Three Conversational Contexts.

	Professional	**Neutral**	**Flirtatious**
Mean pitch	198.45 (19.84)$_a$	202.04 (18.97)$_b$	201.11 (18.27)$_{a,b}$
Median pitch	191.99 (20.52)$_a$	194.25 (19.64)$_a$	194.21 (19.00)$_a$
Minimum pitch	97.88 (5.08)$_a$	99.27 (6.33)$_a$	97.13 (7.59)$_a$
Maximum pitch	373.55 (34.39)$_a$	377.11 (28.95)$_a$	373.03 (38.81)$_a$

Note. Means (*M*) and standard deviations (*SD*) given for each pitch parameter in each context.

[a] Within a row, cells that do not share a subscript are significantly different from each other at *p* < .05.

Despite the partial support of the hypothesis demonstrated by participants' mean pitch, there were no significant findings pertaining to the other three dependent variables, all $F(2,60)$'s < 2.40, all *p*'s > .100. See Table 2.

Discussion

In line with the findings of Sorokowski et al. (2019) and Leongómez et al. (2017), participants in Study 2 spoke with a lower mean pitch in the professional context than in the neutral context. However, Study 2 produced no other significant findings. One of the most salient issues that arose during data collection for Study 2 was the presence of background noise. The values for minimum and maximum pitch shown in Table 2 are unnaturally low and high, respectively. Before the pitch range was set to 100-400 Hz for analysis, the Praat pitch analysis regularly found maximum pitches in the range of 490-500 Hz and minimum pitches in the range of 60-70 Hz, both of which are far outside the typical range of female speaking voices. Even accounting for the possibility of a participant occasionally producing an extremely low vocal fry or a high-pitched exclamation, it would not be the case that every participant's response to every question would have such a large pitch range. Study 2 was conducted in an academic building whose classrooms did not have acoustic muffling or other soundproofing features. It is likely that the extremely low pitches came from the humming of an air conditioning or heating unit in the classroom. On many occasions, a clanking noise from the unit was also audible, which may have been the source of the extremely high pitches. Attempts to remove extraneous background noise from each audio file were unsuccessful, as the Praat pitch analysis still found the same overly low and high pitches. As such, the data for minimum and maximum pitch likely were wildly inaccurate and did not reflect participants' real pitch ranges.

The fact that the minimum and maximum pitches did not differ significantly across contexts suggests that an outside factor influenced Praat's interpretation of each recording's minimum and maximum pitch. However, participants' mean pitch did fluctuate across conditions, which implies that some pitch changes (at least according to the parameter of mean

pitch) may have been due to context differences, as hypothesized. By contrast, there were no

significant differences in participants' median pitch across contexts. Notably, though, the

difference in participants' median pitch between the professional and neutral contexts was in the

hypothesized direction, albeit not significantly. In fact, all four pitch variables had lower values

in the professional context than in the neutral context, though only the difference in mean pitch

between contexts was statistically significant. While the median, minimum, and maximum pitch

data were not strong enough to officially support the hypothesis, the fact that these three

variables shared a pattern with the significant results found for mean pitch is promising.

In addition to the pitch range issues caused by environmental factors during data

collection, the study design may also have contributed to the lack of significant findings. While

studies like Pisanski et al. (2018) examined participants' pitch modulation in an explicitly

flirtatious context—a speed-dating scenario—the "flirtatious" context in Study 2 was

considerably more neutral than that in Pisanski et al. The photo chosen to represent "Quinn" was

of a conventionally attractive (fit, tall, good-looking) young adult male, and Quinn asked the

participants to describe their perfect first date. However, the context itself was not presented to

participants in a way that would necessarily have primed them for flirtatiousness. Participants

were told that Quinn had been recruited from a nearby university, but the experimenter did not

comment on his supposed attractiveness, say that he was single, or otherwise hint that the

participants' conversations with Quinn were intended to be flirtatious. As such, participants may

not have understood, even unconsciously, that their interaction with Quinn had, or was supposed

to have, a flirtatious tint. Similarly, although ratings of the flirtatiousness of the flirtatious

context exceeded flirtatiousness ratings of the neutral and professional context, the mean rating

of the conversation with Quinn was hardly higher than 3, or "somewhat flirtatious," on the

manipulation check's 5-point Likert scale. Even though the conversation was judged more flirtatious than the other two, it was not particularly flirtatious in its own right. Whether participants were attracted to Quinn at all was also unknown, as they did not report on feelings of attraction during the study. If they were not attracted to him, then they likely would not have attempted to flirt and might not have modulated their pitch.

By contrast, the professional context was far more effective in coming across as professional. Participants' mean rating of the professionalism of the conversation with Adam was nearly a 5, or "very professional," on the manipulation check's 5-point Likert scale. This exceptionally high rating was likely due to the format of the context itself. Studies in the field of pitch modulation, such as Leongómez et al. (2017), have also utilized a simulated interview to elicit professional speech from participants. In Study 2, the participants were informed ahead of time that they would be asked interview questions, which may have primed them to act more professionally and could explain why their pitch was significantly lower in the professional context than in the neutral context. The same type of setup was not employed in the flirtatious and neutral contexts. If it had been, the results of Study 2 might have been more in line with the research hypotheses.

General Discussion

The studies presented here have contributed to the field's understanding of pitch-based impression formation and pitch modulation by expanding on prior research and attempting to fill several gaps in the literature. Specifically, Study 1 was the first to examine the relationship between pitch and perceptions of speaker sociability and intelligence. The latter two variables had not yet been studied for any gender's voice pitch. Additionally, Study 1 is among the first to have investigated the relationship between women's pitch and perceptions of speaker confidence

and persuasiveness. Although confidence and persuasiveness had been studied in relation to

men's voices, far less research had been dedicated to women's voice pitch before Study 1.

Though not all of the hypotheses for Study 1 were fully confirmed, the results nonetheless

demonstrate that pitch plays a role in impression formation and that, mostly in line with the

study's predictions, traditionally masculine traits are more closely associated with lower-pitched

female voices and vice versa.

Study 2 also made meaningful contributions to the literature on pitch modulation. All

previous research (Fraccaro et al., 2011; Hughes et al., 2010; Leongómez et al., 2017; Pisanski et

al., 2018; Sorokowski et al., 2019) in the field had only compared pitch in one experimental

context (i.e., flirtatious) to a baseline (i.e., neutral), but not to another experimental context.

Study 2, by contrast, directly compared participants' pitch in two experimental contexts

(flirtatious and professional) to each other as well as to a baseline neutral context. Additionally,

Study 2 was the first to examine pitch modulation in conversations that took place over a video-

conferencing app (Zoom) rather than in person or via a hypothetical voicemail procedure. The

ongoing COVID-19 pandemic has significantly increased the use of Zoom and other video-

conferencing apps. It is therefore important to account for the ubiquity of virtual communication

systems when studying a phenomenon relating to interpersonal interaction, such as pitch

modulation. Though the bulk of Study 2's predictions were not confirmed by the data (due in

large part to the technical difficulties surrounding measurements of minimum and maximum

pitch), the results indicated that participants lowered their mean pitch in the professional context,

as hypothesized. Furthermore, the data for minimum, maximum, and median pitch all showed

the same (nonsignificant) difference between the professional and neutral contexts, which may

suggest that a trend does exist and could be uncovered by a future study.

Despite addressing several gaps in the literature, the research presented here was limited in scope in a major way: its concept of womanhood. Both studies examined the voices of cisgender women specifically, and the population of interest in Study 2 was further limited to cisgender women attracted to men. As described earlier, this subset of the population was chosen because the vocal cords of people who were assigned female at birth are anatomically distinct from the vocal cords of people who were assigned male at birth. Furthermore, Study 2 examined people who identify as female in addition to having female vocal cords, as people who identify as female may conduct themselves differently—including how they present themselves in various conversational contexts—than people who identify with other genders. It was also assumed that making attraction to men one of the eligibility criteria would assist with examining pitch modulation towards the male target in the flirtatious context. While each of these criteria had a purpose relative to the research question, they did limit the sample and its generalizability to the greater population. Indeed, most, if not all, pitch research has thus far been limited to examining the voices of cisgender men and women. No pitch modulation research, for example, has investigated transgender people's voices cross-contextually. Transgender individuals may or may not feel the same inclination to lower their pitch cross-contextually, or they may feel even stronger pressure to do so. Trans men and trans women may perform pitch modulation in different ways, which in turn may be different from the approach of people who identify with other genders. Future research will need to include the voices of people whose gender does not match what they were assigned at birth.

Furthermore, the literature on pitch and impression formation has not examined the possibility of interactions between pitch and other traits, such as outward presentation, during impression formation. How might people react to a woman with a low voice who dresses in a

highly feminine style? Would she still be considered as competent as she might have been if the listener had only heard her voice? In Study 1, participants listened to female voices but were told nothing else about the speakers. This methodology—voices in isolation—is common in the field. However, it fails to consider the fact that in most interactions, people forming impressions would have access to considerably more information (including visual and behavioral cues) than just pitch. In these scenarios, pitch may still play a significant role in impression formation, but it could be a less salient part of the process than if it were the only available cue. In Study 2, as well, participants knew that their conversation partners could only hear them, so they may have put more energy into modulating their pitch than they would have otherwise. Had they been seen by their partners during the conversations, they may have monitored their body language or other visual cues more than their pitch. Future studies should examine the potential relationships between visual and auditory cues, which have gone largely understudied.

Its limitations aside, the primary goal of the research presented here was to understand the ways in which women's pitch interacts with both impression formation and impression management. Study 1 showed the importance of pitch as a factor when forming impressions of a woman's competence, confidence, and intelligence, among other traits. Women in the real world might be considered more or less professional based on their voices, because in male-dominated fields, professionalism is often equated with characteristics associated with masculinity, such as competence. Yet this vision of professionalism is inherently unequal. If the expectation of an ideal candidate for a certain field is one whose voice indexes characteristics like competence and confidence, which are more strongly associated with masculinity than femininity, women begin at a disadvantage. The vast majority of men have lower voices than even the lowest-pitched women (Titze, 2000), and thus can more easily index masculinity and masculinity-linked traits.

Women with naturally sex-atypical (i.e., low) voices might index such traits as well, but women with high voices will encounter more obstacles to being taken seriously. It is little wonder that so many sectors lack gender equality. With the expectation that a candidate will project confidence, competence, intelligence, and many more traits, women whose voices do not naturally index these traits cannot break into male-dominated fields.

Women who are aware of the influence of pitch on impression formation may consciously or unconsciously utilize pitch to their advantage. Given the reasons described above, women may feel the need to project a professional image in contexts like job interviews and performance reviews. In line with these considerations, Study 2 demonstrated that women do lower their pitch in professional contexts. Whether participants were aware or not that they were manipulating the way they were perceived, they attempted on some level to change their self-presentation in an interview context. Interpreting these results, and those from Study 1, as a sign that women can and should use pitch to their advantage would be overly simplistic. To discuss only the thought process of women who lower their pitch to appear more masculine would ignore the root of the problem: why should women have to tamp down their femininity to be given serious consideration as job candidates? Participants in Study 2, and women in the workforce in general, could choose to disguise their true pitch in interviews to increase their chances of entering a male-dominated career, but they cannot be expected to maintain this charade forever. With luck, some employers might maintain their initial impressions of female hires who disguise their voices in interviews, but even this best-case scenario is hardly a solution to the broader systemic issue at hand.

What actions can be taken, then, to resolve the issues that make women feel that they need to come across as masculine to be taken seriously? Such a reworking of societal inequality

is beyond the scope of this paper, but one place to start might be an early intervention so that children understand that any career is open to both men and women. Cartei et al. (2021) observed that children as young as nine years old gave high-pitched female voices the lowest competence ratings for stereotypically male jobs and the highest competence ratings for stereotypically female jobs. It seems that even elementary-school-aged children may have already formed judgments about the types of jobs that are appropriate for women and men. Perhaps an intervention to unpack these biases could be a first step in creating a more equitable society for women seeking male-dominated careers. However, the studies presented here have demonstrated two things: women's voices are indeed used as cues in impression formation (and not always to the benefit of the speaker), and whether or not women are aware of this phenomenon, they may still feel the need to change their self-presentation via pitch to appear more masculine in professional contexts.

References

Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of Voice, 13*(3), 424-446. https://doi.org/10.1016/S0892-1997(99)80048-4

Anderson, N. H. (1962). Application of an additive model to impression formation. *Science*, *138*(3542), 817-818. https://doi.org/10.1126/science.138.3542.817

Ansorge, U. (2010). Priming. In E. B. Goldstein (Ed.), *Encyclopedia of perception* (Vol. 1, pp. 825-827). SAGE Publications, Inc., https://dx.doi.org/10.4135/9781412972000.n264

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, *37*(5), 715–727. https://doi.org/10.1037/0022-3514.37.5.715

Aşiret, G.D., & Kapucu, S. (2016). The effect of reminiscence therapy on cognition, depression, and activities of daily living for patients with Alzheimer disease. *Journal of Geriatric Psychiatry and Neurology*, *29*(1), 31–37. https://doi.org/10.1177/0891988715598233

Banai, P. I. (2017). Voice in different phases of menstrual cycle among naturally cycling women and users of hormonal contraceptives. *PloS one*, *12*(8), 1-13. https://doi.org/10.1371/journal.pone.0183462

Banaji, M. R., Hardin, C., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. *Journal of Personality and Social Psychology*, *65*(2), 272–281. https://doi.org/10.1037/0022-3514.65.2.272

Black, D. E., & Vance, M. D. (2021). Do first impressions last? The impact of initial assessments and subsequent performance on promotion decisions. *Management Science*, *67*(7), 4556–4576. https://doi.org/10.1287/mnsc.2020.3703

Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer (Version 6.0.28).

http://www.praat.org/

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance

and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59.

https://doi.org/10.1016/j.anbehav.2011.03.024

Bryant, F. B., Smart, C. M., & King, S. P. (2005). Using the past to enhance the present:

Boosting happiness through positive reminiscence. *Journal of Happiness Studies*, *6*(3),

227–260. https://doi.org/10.1007/s10902-005-3889-4

Buenrostro, P. (2018). [Photograph of man in black suit jacket]. Unsplash.

https://unsplash.com/photos/CSAEI7OGubA

Carmen, R. A., Geher, G., Glass, D. J., Guitar, A. E., Grandis, T. L., Johnsen, L., Philip, M. M.,

Newmark, R. L., Trouton, G. T., & Tauber, B. R. (2013). Evolution integrated across all

islands of the human behavioral archipelago: All psychology as evolutionary psychology.

EvoS Journal: The Journal of the Evolutionary Studies Consortium, *5*(1), 108–126.

Cartei, V., Oakhill, J., Garnham, A., Banerjee, R., & Reby, D. (2021). Voice cues influence

children's assessment of adults' occupational competence. *Journal of Nonverbal

Behavior*, *45*(2), 281–296. https://doi.org/10.1007/s10919-020-00354-y

del Engelsen, B. (2018). [Photograph of man wearing black and red pullover hoodie]. Unsplash.

https://unsplash.com/photos/eNyXK17so6A

Digirolamo, G. J., & Hintzman, D. L. (1997). First impressions are lasting impressions: A

primacy effect in memory for repetitions. *Psychonomic Bulletin & Review*, *4*(1), 121–

124. https://doi.org/10.3758/bf03210784

Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, *12*(4), 453–476. https://doi.org/10.1111/j.1467-9841.2008.00374.x

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

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. https://doi.org/10.1068/p5514

Fraccaro, P. J., Jones, B. C., Vukovic, J., Smith, F. G., Watkins, C. D., Feinberg, D. R., Little, A. C., & Debruine, L. M. (2011). Experimental evidence that women speak in a higher voice pitch to men they find attractive. *Journal of Evolutionary Psychology*, *9*(1), 57–67. https://doi.org/10.1556/jep.9.2011.33.1

Fraccaro, P. J., O'Connor, J. J. M., Re, D. E., Jones, B. C., DeBruine, L. M., & Feinberg, D. R. (2013). Faking it: Deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, *85*(1), 127–136. https://doi.org/10.1016/j.anbehav.2012.10.016

Goffman, E. (1959). *The presentation of self in everyday life*. Anchor Books.

Griffith, E., & Woo, E. (2022, March 22). Elizabeth Holmes found guilty of four charges of fraud. *The New York Times*. https://www.nytimes.com/live/2022/01/03/technology/elizabeth-holmes-trial-verdict#elizabeth-holmes-conviction-appeal

Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2018). Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion.

Personality and Social Psychology Bulletin, *45*(3), 389–405.

https://doi.org/10.1177/0146167218787805

Hallemann, C. (2019, March 22). Is Elizabeth Holmes's deep voice just another part of her scam?

Town & Country. https://www.townandcountrymag.com/society/money-and-

power/a26855364/elizabeth-holmes-deep-voice/

Hari Kumar, K. V., Garg, A., Ajai Chandra, N. S., Singh, S. P., & Datta, R. (2016). Voice and

endocrinology. *Indian Journal of Endocrinology and Metabolism*, *20*(5), 590–594.

https://doi.org/10.4103/2230-8210.190523

Herper, M. (2016, June 1). *From $4.5 billion to nothing: Forbes revises estimated net worth of*

Theranos founder Elizabeth Holmes. Forbes.

https://www.forbes.com/sites/matthewherper/2016/06/01/from-4-5-billion-to-nothing-

forbes-revises-estimated-net-worth-of-theranos-founder-elizabeth-

holmes/?sh=829c32636331

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression

formation. *Journal of Experimental Social Psychology*, *13*(2), 141–154.

https://doi.org/10.1016/S0022-1031(77)80007-3

Hughes, S. M., Farley, S. D., & Rhodes, B. C. (2010). Vocal and physiological changes in

response to the physical attractiveness of conversational partners. *Journal of Nonverbal*

Behavior, *34*(3), 155–167. https://doi.org/10.1007/s10919-010-0087-9

Jasieńska, G., Ziomkiewicz, A., Ellison, P. T., Lipson, S. F., & Thune, I. (2004). Large breasts

and narrow waists indicate high reproductive potential in women. *Proceedings of the*

Royal Society of London, *271*, 1213–1217. https://doi.org/10.1098/rspb.2004.2712

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*, *79*(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Krahé, B., Uhlmann, A., & Herzberg, M. (2021). The voice gives it away. *Social Psychology*, *52*(2), 101–113. https://doi.org/10.1027/1864-9335/a000441

Leongómez, J. D., Mileva, V. R., Little, A. C., & Roberts, S. C. (2017). Perceived differences in social status between speaker and listener affect the speaker's vocal characteristics. *PLOS ONE*, *12*(6), 1–21. https://doi.org/10.1371/journal.pone.0179407

Nate. (2018). [Photograph of man wearing blue shirt]. Unsplash. https://unsplash.com/photos/J5U-22o1ubw

O'Connor, J. J. M., Re, D. E., & Feinberg, D. R. (2011). Voice pitch influences perceptions of sexual infidelity. *Evolutionary Psychology*, *9*(1), 64-78. https://doi.org/10.1177/147470491100900109

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2016). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, *24*(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

Pisanski, K., Oleszkiewicz, A., Plachetka, J., Gmiterek, M., & Reby, D. (2018). Voice pitch modulation in human mate choice. *Proceedings of the Royal Society B: Biological Sciences*, *285*, 1–8. https://doi.org/10.1098/rspb.2018.1634

Schild, C., Stern, J., & Zettler, I. (2019). Linking men's voice pitch to actual and perceived trustworthiness across domains. *Behavioral Ecology*, *31*(1), 164–175. https://doi.org/10.1093/beheco/arz173

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, *23*(3-4), 193–229. https://doi.org/10.1016/s0271-5309(03)00013-2

Singh, D., Dixson, B. J., Jessop, T. S., Morgan, B., & Dixson, A. F. (2010). Cross-cultural consensus for waist–hip ratio and women's attractiveness. *Evolution and Human Behavior*, *31*(3), 176–181. https://doi.org/10.1016/j.evolhumbehav.2009.09.001

Sorokowski, P., Puts, D., Johnson, J., Żółkiewicz, O., Oleszkiewicz, A., Sorokowska, A., Kowal, M., Borkowska, B., & Pisanski, K. (2019). Voice of authority: Professionals lower their vocal frequencies when giving expert advice. *Journal of Nonverbal Behavior*, *43*(2), 257–269. https://doi.org/10.1007/s10919-019-00307-0

Srull, T. K., & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, *37*(10), 1660–1672. https://doi.org/10.1037/0022-3514.37.10.1660

Stern, J., Schild, C., Jones, B. C., DeBruine, L. M., Hahn, A., Puts, D. A., Zettler, I., Kordsmeyer, T. L., Feinberg, D., Zamfir, D., Penke, L., & Arslan, R. C. (2021). Do voices carry valid information about a speaker's personality? *Journal of Research in Personality*, *92*, 1–14. https://doi.org/10.1016/j.jrp.2021.104092

Streeter, L. A., Krauss, R. M., Geller, V., Olson, C., & Apple, W. (1977). Pitch changes during attempted deception. *Journal of Personality and Social Psychology*, *35*(5), 345–350. https://doi.org/10.1037/0022-3514.35.5.345

Taylor, S. E. (1981). "The interface of cognitive and social psychology". In J. H. Harvey (Ed.), *Cognition, social behavior, and the environment* (pp. 189-211). L. Erlbaum.

Titze, I. R. (2000). *Principles of voice production*. National Center for Voice and Speech.

Zheng, Y., Compton, B. J., Heyman, G. D., & Jiang, Z. (2020). Vocal attractiveness and

   voluntarily pitch-shifted voices. *Evolution and Human Behavior*, *41*(2), 170–175.

   https://doi.org/10.1016/j.evolhumbehav.2020.01.002