Investigating the Effects of Caffeine on Phonation

Elizabeth Erickson-Levendoski and Mahalakshmi Sivasankar, West Lafayette, Indiana

Summary: Objective. A core component of vocal hygiene programs is the avoidance of agents that may dry the vocal folds. Clinicians commonly recommend that individuals reduce caffeine intake because of its presumed dehydrating effects on the voice. However, there is little evidence that ingestion of caffeine is detrimental to voice production. The first objective of this study was to evaluate whether caffeine adversely affects voice production. The second objective was to evaluate if caffeine exacerbates the adverse phonatory effects of vocal loading.

Study Design. Prospective, double-blinded, sham-controlled study.

Methods. Sixteen healthy adults participated in two sessions where they consumed caffeine (caffeine concentration = 480 mg) or sham (caffeine concentration = 24 mg) beverages. Voice measures (phonation threshold pressure and perceived phonatory effort) were collected. Subjects then completed a vocal loading challenge and voice measures were obtained again.

Results. There were no significant differences in voice measures between the caffeine and sham conditions. Ingestion of caffeine did not adversely affect voice production (P > 0.05) or exacerbate the detrimental phonatory effects of vocal loading (P > 0.05).

Conclusions. Our findings contribute to emerging knowledge on the effects of caffeine on voice production. Recommendations to completely eliminate caffeine from the diet, as a component of a vocal hygiene program, should be evaluated on an individual basis.

Key Words: Caffeine–Voice production–Dehydration.

INTRODUCTION
Avoiding agents that dehydrate the vocal folds is an integral component of vocal hygiene education. Systemic dehydration is detrimental to voice production. Because of the presumed systemic drying action of caffeine, voice clinicians recommend avoidance of caffeinated beverages. This remains common clinical adage even with little supporting evidence that caffeine consumption induces negative changes to voice. Consequently, the first objective of this study was to investigate whether caffeine is detrimental to phonation. Voice measures were compared with subjects, who each consumed beverages that contained large amounts of caffeine (caffeinated beverages) and negligible amount of caffeine (sham beverages). Voice measures included phonation threshold pressure (PTP) and perceived phonatory effort (PPE). PTP is an appropriate voice measure to evaluate the phonatory effects of caffeine because PTP estimates the lung pressure that is required for maintaining voicing at threshold loudness and is sensitive to systemic dehydration. Another useful indicator of the phonatory effects of caffeine may be participant ratings of effort required for voice production (PPE). To determine if caffeine changes voice, PTP and PPE ratings were obtained after consuming caffeinated and sham beverages. An increase in PTP and/or PPE following caffeine but not sham beverages would suggest caffeine is detrimental to voice production.

The second objective of this investigation was to evaluate whether caffeine exacerbates the detrimental phonatory effects of vocal loading. This question is clinically relevant because individuals often work and socialize in environments where they engage in vocal loading activity such as prolonged speaking in background noise. It is well known that vocal loading alone adversely affects the voice, and there is emerging evidence that systemic dehydration may exacerbate the negative phonatory effects of vocal loading. To examine the relationship between dehydration and vocal loading in the context of the purported systemic drying action of caffeine, we examined the interaction between caffeine consumption and vocal loading on PTP and PPE. Findings from this investigation are important for the development of evidence-based recommendations regarding the effects of caffeine consumption on voice production.

METHODS
Participants
All procedures used here were approved by the Purdue University Institutional Review Board. Sixteen healthy adults (eight males and eight females; mean age = 23 years) with perceptually normal speech and voice participated in this investigation. Participants reported a negative history for hearing problems, respiratory disease, smoking, refluxing, and prescription medication (except oral contraceptives). All female subjects participated during the follicular phase of their menstrual cycle. Eight individuals had received vocal training (Table 1), either through group and/or solo singing (seven subjects) or theatrical speaking (1 subject). Trained subjects were purposely included because they frequently engage in vocal loading-type activities and are also instructed to avoid caffeinated beverages to maintain a healthy voice.
TABLE 1.
Subject Characteristics

<table>
<thead>
<tr>
<th>Subject Number</th>
<th>Gender</th>
<th>Age (Years)</th>
<th>Training (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Male</td>
<td>18</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Male</td>
<td>21</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Male</td>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Male</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>Male</td>
<td>26</td>
<td>None</td>
</tr>
<tr>
<td>6</td>
<td>Male</td>
<td>27</td>
<td>None</td>
</tr>
<tr>
<td>7</td>
<td>Male</td>
<td>20</td>
<td>None</td>
</tr>
<tr>
<td>8</td>
<td>Male</td>
<td>25</td>
<td>None</td>
</tr>
<tr>
<td>9</td>
<td>Female</td>
<td>24</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>Female</td>
<td>20</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>Female</td>
<td>27</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>Female</td>
<td>32</td>
<td>10</td>
</tr>
<tr>
<td>13</td>
<td>Female</td>
<td>19</td>
<td>None</td>
</tr>
<tr>
<td>14</td>
<td>Female</td>
<td>20</td>
<td>None</td>
</tr>
<tr>
<td>15</td>
<td>Female</td>
<td>20</td>
<td>None</td>
</tr>
<tr>
<td>16</td>
<td>Female</td>
<td>23</td>
<td>None</td>
</tr>
</tbody>
</table>

Procedure

This investigation utilized a double-blinded, sham-controlled, repeated-measures design. Participants were blinded to the objectives of this study and consented to an investigation examining the “effects of taste and temperature of ingested beverages on voice production.” Likewise, the investigators collecting and analyzing the data were blinded to whether the consumed beverages were caffeinated or sham. Each subject participated in two experimental sessions scheduled at the same time on two consecutive days. In each session, subjects ingested either caffeinated beverages (caffeine content ~480 mg) or sham beverages (caffeine content ~24 mg). The sessions were counterbalanced across subjects. The caffeinated beverages consisted of two 12-ounce caffeinated Starbucks coffees, whereas the sham beverages consisted of two 12-ounce Starbucks decaffeinated coffees. During both the caffeine and sham sessions, participants left the laboratory following consumption of the first beverage and returned for the second beverage and voice testing within 2.5–3 hours. Total duration of each experimental session was 5 hours beginning with consumption of the first beverage. This timeline is within the elimination half-life of caffeine. Participants were instructed not to consume caffeine before the experimental sessions. Only one participant reportedly consumed a 20-ounce soda (caffeine content ~63 mg) before the caffeinated session.

Data collection procedures were identical for both sessions and completed in a humidity-controlled environment (70% ± 6%). Following consumption of the second beverage, voice measures were collected. These voice measures are hereafter referred to as preloading measures. Participants then completed a vocal loading challenge. Specifically, the vocal loading challenge for the nontrained subjects, and the one subject trained in theatrical speaking consisted of 35 minutes of reading aloud in background noise (cafeteria noise: 65 dB sound pressure level). The vocal loading challenge for the subjects trained in group and/or solo singing consisted of 35 minutes of singing in identical background noise. Voice measures were reassessed and this time point is hereafter referred to as Loading35. Subjects then completed an additional 35 minute vocal loading challenge as described above, and voice measures were reobtained, and this time point is hereafter referred to as Loading70. The voice measures included PTP and PPE.

Initially, subjects were trained on the PTP task. The training consisted of subjects repeatedly practicing the PTP task (enunciating a five to seven /p/ syllable string in a soft voice, on a single breath at a rate of 1.5 syllables/second guided with a metronome). Participants were provided visual feedback and investigator modeling during the training. Participants were trained on the PTP task at two different pitches. These pitches were the 10th and 80th percent pitch of each subjects’ maximum pitch range. Extreme pitches were purposely selected based on prior research that the effects of systemic dehydration are most apparent at PTP productions at high pitch. Participants were deemed trained when the following criteria were met at both the 10th and 80th percent of the pitch range: /p/ pressure peaks of equal height, consistent quiet voicing, and oral flows approaching 0 mL/s with nostrils occluded. Once trained to the above criteria, PTP measures were collected at the 10th and 80th pitches (hereafter referred to as PTP10 and PTP80, respectively). At least five repetitions of the PTP task were collected at PTP10 and PTP80. The instrumentation for PTP collection included a circumferentially vented pneumotachograph face mask coupled to transducers (Glottal Enterprises, Syracuse, NY) for the measurement of oral flow (sampling rate = 1000 Hz) and oral pressure (sampling rate = 100 Hz). Within each /p/ syllable string, the middle three /p/ peaks were selected, and adjacent peaks were measured to estimate lung pressure during the intervening /i/. Estimated lung pressures from each string were averaged for the final PTP at each pitch. PTP analysis was completed by investigators blinded to the identity of the experimental session. To ensure PTP measurement accuracy between investigators, 10% of the /p/ syllable strings were remeasured, and intrarater reliability was computed. To ensure PTP measurement accuracy within an investigator, 10% of the /p/ syllable strings were remeasured and intrarater reliability was also computed. Inter- (r = 0.98) and intrarater reliability (r = 0.99) revealed strong correlations.

To obtain PPE ratings, subjects sang “Happy Birthday” in a soft voice starting at the 50th percent pitch of the maximum phonation range. Subjects rated the effort for singing on a 10” visual analog scale with anchors of “no vocal effort” and “maximum vocal effort.” Subjects rated PPE at three time points (Preloading, Loading35, and Loading70). Analysis of PPE was completed by measuring the distance in inches, from the “no vocal effort” anchor to the subject’s PPE rating, by investigators blinded to the identity of the experimental session. To ensure PPE measurement accuracy between investigators, 10% of PPE ratings were remeasured and intrarater reliability was computed. To ensure PPE measurement accuracy within an investigator, 10% of PPE ratings were remeasured and intrarater reliability was computed. Inter- (r = 0.99) and intrarater reliability (r = 0.99) revealed strong correlations.
Statistical analysis
Data were analyzed using a mixed general linear model analysis of variance (ANOVA; alpha level = .05). The t tests were used for post hoc analyses. The independent variables were condition (caffeine/sham) and vocal loading (Preloading, Loading35, and Loading70). The dependent variables were PTP10, PTP80, and PPE. Because the core objectives of this study were to investigate the effects of caffeine on phonation and the interaction between vocal loading and caffeine, and because visual inspection of the data did not reveal group differences because of vocal training, data for trained and nontrained subjects were merged in statistical analysis.

RESULTS
This investigation evaluated the effects of caffeine and vocal loading on voice production. Caffeine consumption did not adversely affect voice production (Figures 1 and 2). There were no significant differences between the caffeine and sham conditions for PTP10, PTP80, or PPE (P > 0.05). Ingestion of caffeine or sham beverages did not worsen the effects of vocal loading for PTP10, PTP80, or PPE (P > 0.05).

However vocal loading, alone, significantly increased PTP10 and PTP80 (P < 0.01, Figure 1). Post hoc testing revealed that PTP10 and PTP80 significantly increased after 35 minutes of vocal loading and after 70 minutes of vocal loading (P < 0.01, Figure 1). Conversely, PPE did not increase with vocal loading (P > 0.05, Figure 2). Details from the statistical analyses are presented in Table 2.

DISCUSSION
Over the past 20 years, caffeine consumption has increased substantially.17 Caffeine is the world’s most widely used dietary stimulant18 with the average adult consuming approximately 227 mg caffeine per day17,19 in its various forms. Following oral ingestion, caffeine is rapidly absorbed by the gastrointestinal tract and elicits numerous physiological and biochemical effects13,20 For example, caffeine consumption increases excretion of water and electrolytes.13 This diuretic action of caffeine is presumed to induce systemic dehydration.2,5 Given that systemic dehydration negatively affects voice production,4,9 voice clinicians strongly caution against excessive caffeine use1,21,22.

To the best of our knowledge, only one pilot investigation has examined the effects of caffeine consumption (250 mg caffeine tablets) on voice.5 The results of this pilot investigation were nonsignificant, but no definite conclusions could be drawn secondary to small sample size and the lack of the following: double blinding, within-subject controls, and temporal controls. The present study incorporated a double-blinded, sham-controlled, repeated-measures design, where each participant consumed both caffeinated and sham beverages in counterbalanced order, on different days, to investigate whether caffeine consumption is detrimental to voice production. Although the caffeine concentrations utilized here are greater than that consumed by the average adult,17 this dosage was purposely chosen. Past research suggests that the diuretic response to caffeine occurs at doses exceeding 300 mg.23 It is of note that despite the high dose of
caffeine intake in this study (480 mg) no changes in PTP and PPE were observed.

The absence of a significant effect in PTP and PPE following consumption of caffeinated beverages is surprising because these measures are thought to be sensitive to changes in vocal fold hydration status.\(^4,9\) It is possible that caffeine does not produce a diuretic effect of sufficient magnitude to elicit negative voice changes in healthy subjects who maintain a regular balanced diet.\(^20\) Although the extent of fluid loss post-caffeine ingestion was not quantified here, previous voice research demonstrates that adverse voice changes (as measured by PTP) occur only after substantial systemic fluid loss induced by dialysis\(^9\) or the loop diuretic, Lasix.\(^4\) The effects of caffeine intake on systemic physiology and biochemistry may provide further insight into the nonsignificant changes in PTP and PPE observed in this study. Healthy individuals quickly develop a tolerance to the diuretic effects of caffeine, a tolerance strengthened by regular caffeine consumption.\(^20,23\) In this investigation, all but two participants reported using caffeine, although the frequency of caffeine use ranged from once per month to twice per day. Further study is required on how tolerance to caffeine affects voice production. Future studies should also include a caffeine withdrawal period, because the diuretic action of caffeine is most noted in individuals deprived of caffeine for a period of days.\(^24\) We did not impose a withdrawal period in this study to mimic realistic conditions for participants.

Systemic dehydration has been previously shown to exacerbate the negative effects of vocal loading.\(^3\) In the current investigation, ingestion of caffeine did not have a significant effect on voice production in either trained or nontrained speakers exposed to vocal loading. One possible interpretation of this finding is that the dose of caffeine used here did not induce the extent of systemic dehydration needed to exacerbate the ill effects of vocal loading. However, the vocal loading task, alone, significantly increased PTP independent of caffeine consumption. Therefore, it is also possible that the detrimental effects of vocal loading were of sufficient magnitude to override the negative phonatory effects, if any, associated with caffeine intake. Surprisingly, PPE did not increase following vocal loading. The rationale for this finding is unclear, but could be related to the nature of the singing task used to elicit PPE ratings.

CONCLUSIONS

Findings from the present study suggest that a high dose of caffeine does not adversely affect PTP or PPE measures within the timeline examined. Additional measures may be more sensitive to caffeine-induced perturbations in vocal fold physiology. We selected PTP and PPE measures as these techniques are commonly used to detect changes in vocal fold function. This study lays the groundwork for further investigations that quantify the specific amount of dehydration induced by caffeine, in larger sample sizes, to develop evidence-based recommendations regarding caffeine consumption and voice.

Acknowledgments

We acknowledge the contributions of Kaleena Cruz to data collection and analysis. We also thank Grace Scott, Juli Hoon, and Dakota Robinson.

REFERENCES


