

Sounds hard: Prosodic features reflect effort level and related affective states during exercise

Aviv Emanuel^{a,b,c,1,*}, Inbal Ravreby^{d,e,1,**}

^a School of Psychological Sciences, Tel-Aviv University, Tel-Aviv, Israel

^b School of Public Health, Tel-Aviv University, Tel-Aviv, Israel

^c Sylvan Adams Sports Institute, Tel Aviv University, Tel-Aviv, Israel

^d Department of Brain Sciences, Weizmann Institute of Science, Rehovot, Israel

^e The Azrieli National Institute for Human Brain Imaging and Research, Rehovot, Israel

ARTICLE INFO

Keywords:

Effort
Resistance exercise
Affect
Prosodic features
Pitch

ABSTRACT

Bouts of exercise have a substantial affective influence, which can impact performance and adherence through training programs. Yet, both the level of effort exertion and affective state during exercise are hard to monitor without the use of questionnaires, which suffer from certain limitations. Here, we examined whether prosodic features, prominent characteristics of human expression, reflect the effort level and its related affect during bouts of exercise. We extracted prosodic features from verbal affective valence ratings recorded in a previously published study ($n = 20$; 10 women; $n_{\text{obs}} = 2428$) of resistance exercises performed by trained participants until task failure. We found that the mean and SD of the pitch predicted effort-related affective valence and proximity to task failure in the two subsets of the data, and in three separate bouts of exercise. These results imply that mean pitch elevation and the decrease of the SD of the pitch during effort exertion may serve as a signal of distress as task difficulty increases. The consistency of the findings across different exercises suggests that the mean and the SD of the pitch may be used to monitor physical effort and affect in various settings and help uncover the nature of physical effort in its different manifestations.

1. Introduction

Physical effort exertion has a substantial affective influence (Ekkekakis et al., 2011; Haines et al., 2020). Such an impact, in turn, can affect well-being and long-term behaviors of both athletes and the general public. For example, affective states during workouts were found to predict burnout (Lemyre et al., 2006) and adherence to exercise programs (Williams et al., 2008). As exercise adherence plays a key role in people's health and well-being (Chen & Wu, 2022; Mandolesi et al., 2018), effectively monitoring exercise related affective responses is crucial.

Effort-related affective response, however, is commonly measured by self-reports via questionnaires, which are difficult to employ during intense exercise (Adams, 2005; Brenner & DeLamater, 2016). Moreover, to date, different questionnaires are used to assess how people feel during bouts of exercise. In some cases, effort perception questionnaires

also measure the affective states during bouts of exercise, and in other cases, questionnaires for specific emotions are employed. Relatedly, one of the reasons that there are competing models of effort perception may be the usage of different questionnaires which relate to different behaviors and subsequently lead to different outcomes (Emanuel et al., 2020; Steele et al., 2016, 2023). An objective and unbiased measure to estimate people's effort level during bouts of exercise and its corresponding affective response is therefore of need, and may serve various scientific and applied fields (De Jong & Elfring, 2010; Dickinson & Villeval, 2008; Yokomizo et al., 2004).

We suggest that prosodic features (i.e., pitch, intensity and speech rate) may serve as a proxy measure for physical effort exertion, as well as its corresponding affective response. Prosodic features are prominent in a wide scope of human-related contexts, and are highly related to how people express themselves (Symons et al., 2016). People use prosodic features to communicate across long distances and are among the most

* Corresponding author. School of Psychological Sciences, School of Public Health and Sylvan Adams Sports Institute, Tel Aviv University, Tel-Aviv, Israel.

** Corresponding author. Department of Brain Sciences and the Azrieli National Institute for Human Brain Imaging and Research, Rehovot, Israel.

E-mail addresses: aviv.emanuel@gmail.com (A. Emanuel), inbalra5991@gmail.com (I. Ravreby).

¹ These authors contributed equally to this work.

prominent characteristics of human expression, reflecting a wide scope of emotional content (Anikin & Lima, 2017; Anikin & Persson, 2017; Anikin & Reby, 2022; Cao et al., 2014; Filippa et al., 2022). For example, prosodic features reflect the evaluations of personality traits (Guidi et al., 2019; Stern et al., 2021) and also emotions such as anger, disgust, fear, happiness and sadness (Cao et al., 2014). In light of the above, the present work suggests that physical effort exertion and its related affect may be reflected in prosodic features. This, in turn, may establish a new framework for studying effort across different contexts.

Here we present an exploratory study in which we examine whether that prosodic features reflect the level of effort exerted and its associated affective valence in various exercises. In order to efficiently monitor effort throughout a task, it is crucial to scale the current performance relative to the maximal ability of each individual. For this purpose, we chose to reanalyze data from a study that included physically effortful tasks (i.e., weight lifting exercises; Emanuel et al., 2020), in which resistance trained participants were asked to reach task failure several times in each session. Participants vocally rated how they felt after each repetition of each exercise, and using their recordings we tested how prosodic features were related to both their affective valence ratings and their proximity to task failure. We chose to focus on the voice intensity, pitch and speech rate, as these were found to be related to emotional expression and impression formation (Cao et al., 2014; De Waele et al., 2019; Gharavian et al., 2010; Koolagudi & Krothapalli, 2011). More specifically, based on the findings that high pitch usually reflects intense emotional experience (Bailen et al., 2019; Cowen & Keltner, 2017; Dietrich et al., 2019), we examined whether the higher the pitch, the closer to task failure people will be and the more negative affect they will express. To the best of our knowledge, this is the first study to examine the relationship between affective state, proximity to task failure and prosodic features in a resistance training setting.

2. Methods

2.1. Materials and procedure

The experimental protocol and participants' demographics are fully described in Emanuel et al. (2020). Briefly, 20 resistance trained participants (10 females; $M_{age} = 30.05$, $SD_{age} = 6.51$) performed four experimental sessions after signing an informed consent form approved by the Institutional Review Board. In the first session, the maximal weight they could lift once was determined (i.e., one repetition-maximum; 1RM) in the bench-press and squat exercises. In the following three sessions, participants completed three sets to task failure with either 70% 1RM bench-press (bench condition), 70% 1RM squat (squat-70% condition), or 80% 1RM squat (squat-80% condition) in a randomized, counterbalanced order. Before and after each set, and after every repetition within the sets, participants verbally reported how they felt on a scale ranging from +5 ("very good") to -5 ("very bad; Hardy & Rejeski, 1989), with their ratings recorded via a clip-on microphone (Boya, BY-M3-OP) attached to the upper side of their shirts. After excluding 87 observations that could not be analyzed due to sound artifacts (e.g., door slam or the experimenter talking in parallel to a participant rating), the sample consisted of 2814 verbal ratings overall for subsequent analyses, of which 2428 observations were within the sets.

2.2. Data pre-processing

For the current analysis, we calculated new prosodic features variables which were not included in the original dataset. To this end, all recordings were exported as mono WAV files and were edited via Audacity (version 2.4.2), to reduce possible noise such as heavy breathing, while keeping the original length of each file intact. This was done manually by setting noise reduction to maximum for each part of the file that did not include an affective valence rating by the participants. Thus,

the resulting recordings included only the affective valence ratings. We then exported the files to PRAAT (Boersma & Weenink, 2021), where the silence threshold was -25 dB, the minimum dip between peaks was 2 dB and the minimum pause duration was 10 ms. We used an in-house script based on de Jong & Wempe (2009) to automatically extract the following aspects of prosodic features: (i) the duration of each affective valence rating; (ii) average, maximum and minimum pitch (i.e., the frequency of the sound), within the boundaries of the expected pitch of 75–500 Hz for men and 120–600 Hz for women (Boersma & Weenink, 2021; Henton, 1995; Manson et al., 2013; Re et al., 2012); (iii) average, maximum and minimum intensity (i.e. how loud the sound was), within the boundaries of the expected human conversation intensity of 55–66 dB (Olsen., 1998); (iv) SD of the pitch and intensity, calculated from windows of 10 ms, was used to estimate voice changes during the speech, for each affective valence rating; (v) the speech rate, estimated by the syllable duration, calculated by the number of syllables in each spoken word, divided by the duration of the affective valence rating. The data and R syntax for the main analyses are available at <https://bit.ly/3jJtGN>.

2.3. Statistical analysis

First, to test whether prosodic features can further contribute to the prediction of proximity to task failure on top of affective ratings alone, we compared a mixed regression model which predicts proximity to task failure from affective valence ratings during exercise, with a model which included the prosodic features on top of the latter predictors. Both models included a random intercept and a random slope for affective valence nested within participants. Importantly, we consider proximity to task failure as a measure of effort rather than fatigue. Effort can be defined as the mobilization of resources in order to carry out instrumental behavior (Sander & Scherer, 2009). Therefore, the level of effort in resistance exercise can be operationalized as the number of repetitions performed in a given time out of the maximal repetition capacity of an individual. In contrast, fatigue can be defined as the decrease in one's maximal capacity to perform a task (Micklewright et al., 2017), and can be operationalized as the reduction in the number of repetitions one is able to perform in a given set when performing it again and again. Therefore, here we assess the link between prosodic features and the level of effort in resistance exercises, rather than the level of fatigue.

Next, to test the robustness and generalizability of our findings, we sought to examine whether any significant effect we find is replicated across two datasets (Fafchamps & Labonne, 2017). To do this, in the main analyses, we divided the original dataset into two subsets of 10 randomly assigned participants. We tested mixed regression models on each subset with all nine prosodic features as predictors and either affective valence ratings or proximity to task failure as outcomes. We added random slopes to the first subset based on the deviance goodness-of-fit index according to the recommendations of Bliese and Ployhart (2002). When adding random slopes significantly improved the model fit in the first subset, we used the same random slopes in the second subset to avoid overfitting. Accordingly, we included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration.

To test replicability across different modalities, we also performed the same analyses described above, after dividing the data of the 20 participants into three exercises – bench-press, squat-70% and squat-80%. In other words, we divided the original dataset into three subsets, one for each exercise, and tested mixed regression models with the nine prosodic features as predictors and either affective valence ratings or proximity to task failure as outcomes. Also in these analyses we included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration.

3. Results

When testing whether adding prosodic features to affective valence ratings can improve the prediction of proximity to task failure, we found that the inclusion of the prosodic features significantly improved the model fit $\chi^2(9) = 97.29, p < 0.001$. This suggests that prosodic features add relevant information about the proximity to task failure that is not transmitted via affective ratings alone. Next, we tested which prosodic features can successfully predict affective valence and proximity to task failure. The results of the mixed regression model predicting affective valence from prosodic features for the two main subsets and for each exercise with standardized (i.e., scaled) predictors are presented in Figs. 1 and 2, respectively. Note that these analyses were for ratings provided within the sets rather than before or after the performance of the sets. The results of the mixed regression model that predicts proximity to task failure from prosodic features for the two main subsets and for each exercise with standardized predictors are presented in Figs. 3 and 4, respectively. We provide tables of the specific statistics for each of the models with unscaled predictors, as well as the prediction of affective valence from prosodic features before and after set completion, in the Supplemental Material section at <https://bit.ly/3jlJtGN> (see Supplemental Tables 1–6; an effect size of pseudo R^2_p for each fixed effect was calculated based on the recommendations of Jaeger et al., 2017). Below we describe the main findings for each dependent variable.

Affective valence. We found that all the pitch characteristics examined – mean, minimum, maximum and SD – predicted affective valence ratings during exercise performance across the two datasets (see Fig. 1 and Table 1). Specifically, we found a negative relationship between the affective state and mean pitch ($p < 0.001$ and $p = 0.009$ in the first and second subset, respectively) as well as between the affective state and maximum pitch ($p < 0.001$ in both subsets). In contrast, we found a positive relationship between the affective state and the pitch minimum ($p < 0.001$ and $p = 0.001$ in the first and second subset, respectively), as

well as between the affective state and the pitch SD ($p < 0.001$ in both subsets). This suggests that pitch characteristics are deeply manifested in affective ratings related to effort exertion. In addition, three intensity characteristics predicted affective valence ratings across the two datasets (see Fig. 1 and Table 1). There was a negative relationship between the affective state and the maximum intensity ($p < 0.001$ in both subsets) and a positive relationship between the affective state and the minimum intensity ($p = 0.023$ and $p = 0.009$ in the first and second subset, respectively). In addition, there was a positive association between the affective state and the intensity SD ($p < 0.001$ and $p = 0.002$ in the first and second subset, respectively). This suggests that intensity characteristics are also manifested in affective ratings related to effort exertion. Last, there was a positive relationship between the speech rate (i.e., syllable duration) and the affective valence ratings in the first dataset ($p = 0.009$) and a trend in the same direction in the second dataset ($p = 0.054$) (see Fig. 1 and Supplemental Table 1). This indicated that the speech rate also expresses affective states during exercise to a certain degree.

When merging the two subsets and testing mixed regression models on each of the three exercises separately, in all three exercises we found that three pitch characteristics predicted affective valence (see Fig. 2 and Table 2). There was a negative relationship between affective valence and the mean pitch ($p = 0.005, p < 0.001$ and $p < 0.001$ in the bench press, squat 70% and squat 80% exercises, respectively) and a negative relationship between the affective valence and maximum pitch ($p = 0.005, p < 0.001$ and $p < 0.001$ in the bench press, squat 70% and squat 80% exercises, respectively). Additionally, we found a positive relationship between affective valence and the pitch SD ($p = 0.035, p < 0.001$ and $p < 0.001$ in the bench press, squat 70% and squat 80% exercises, respectively). This indicates that these pitch characteristics – mean, min and SD – generalize across different exercises in predicting affective valence during physical effort exertion. Regarding intensity characteristics, for each of the three exercises we found a negative

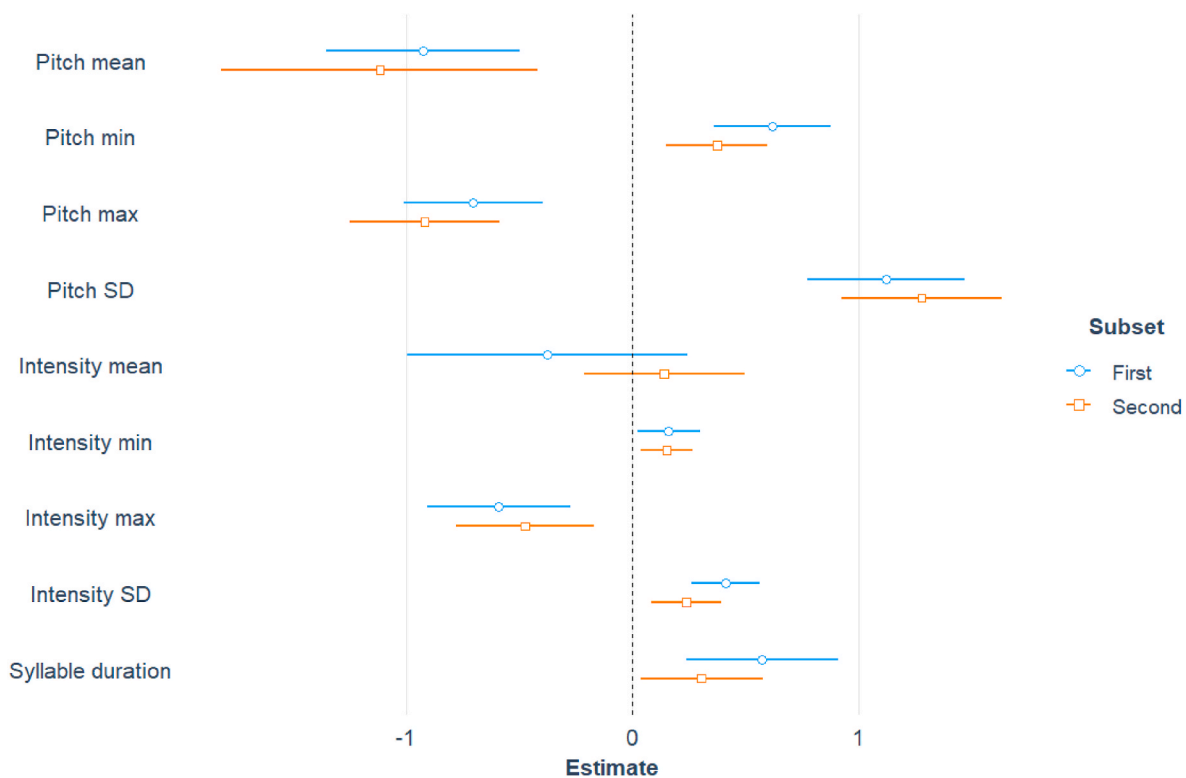


Fig. 1. Mixed model results for the prediction of affective valence by prosodic features as fixed effects for the first subset and the second subset ($n = 10$ each; $n_{\text{obs}} = 1086$ and $n_{\text{obs}} = 1202$, respectively). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration. The circles and squares represent the regression coefficients of the analyses of the first and second subsets, respectively. The error bars represent 95% confidence intervals.

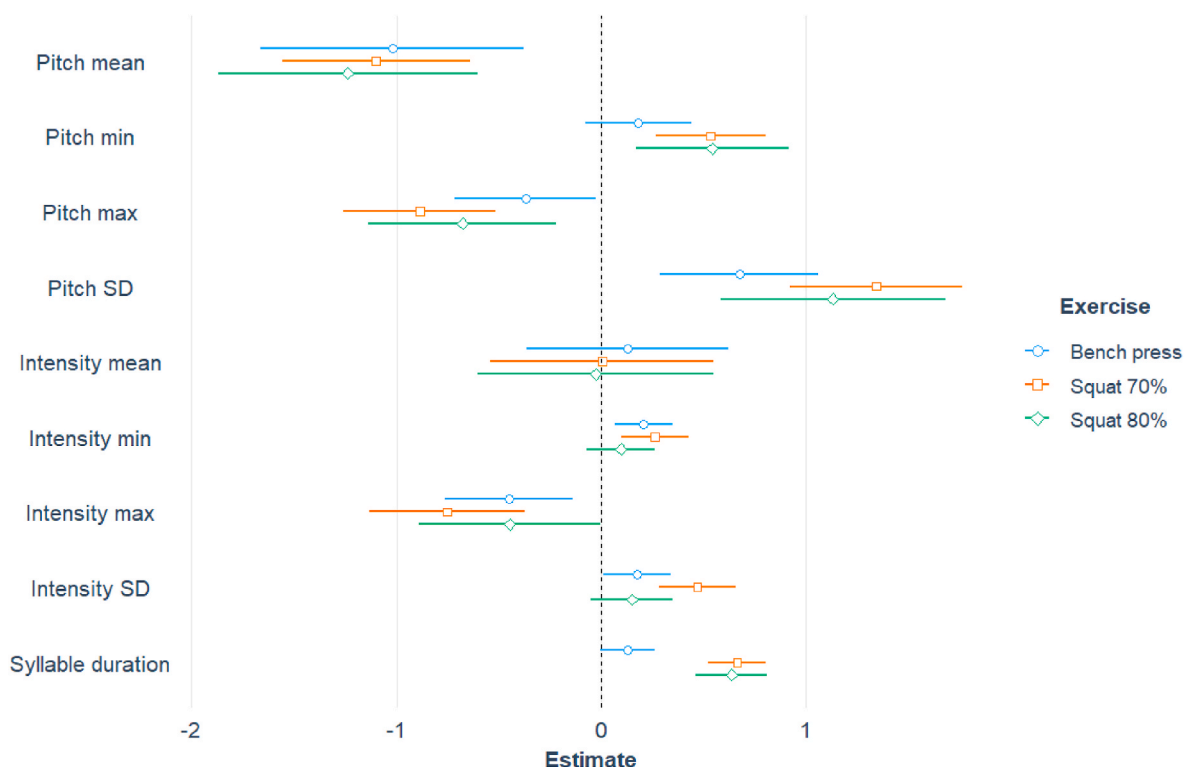


Fig. 2. Mixed model results for the prediction of affective valence by prosodic features as fixed effects for bench press, squat with 70% 1RM and squat with 80% 1RM exercises (bench-press: $n = 19$, $n_{\text{obs}} = 710$; squat 70%: $n = 20$, $n_{\text{obs}} = 971$; squat 80%: $n = 20$, $n_{\text{obs}} = 607$). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration. The circles and squares represent the regression coefficients of the analyses of the first and second subsets, respectively. The error bars represent 95% confidence intervals.

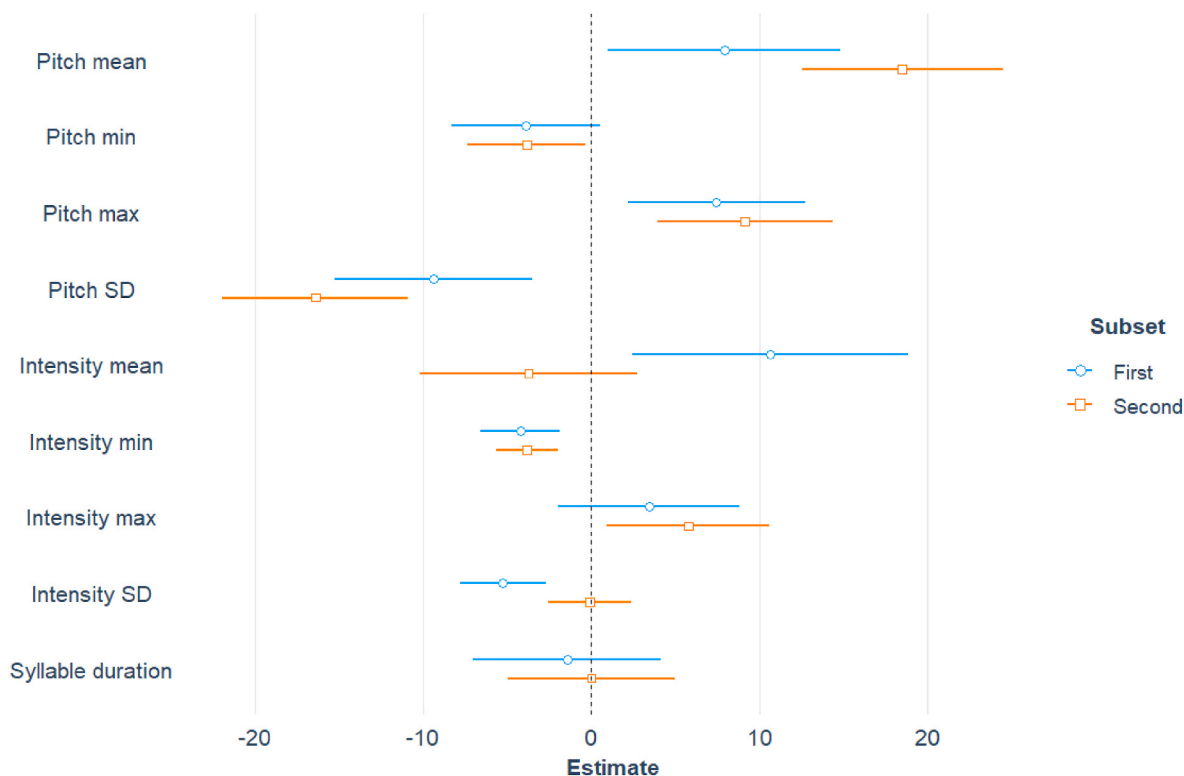


Fig. 3. Mixed model results for the prediction of failure proximity by prosodic features as fixed effects for the first subset and the second subset ($n = 10$ each; $n_{\text{obs}} = 1086$ and $n_{\text{obs}} = 1202$, respectively). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration. The circles and squares represent the regression coefficients of the analyses of the first and second subsets, respectively. The error bars represent 95% confidence intervals.

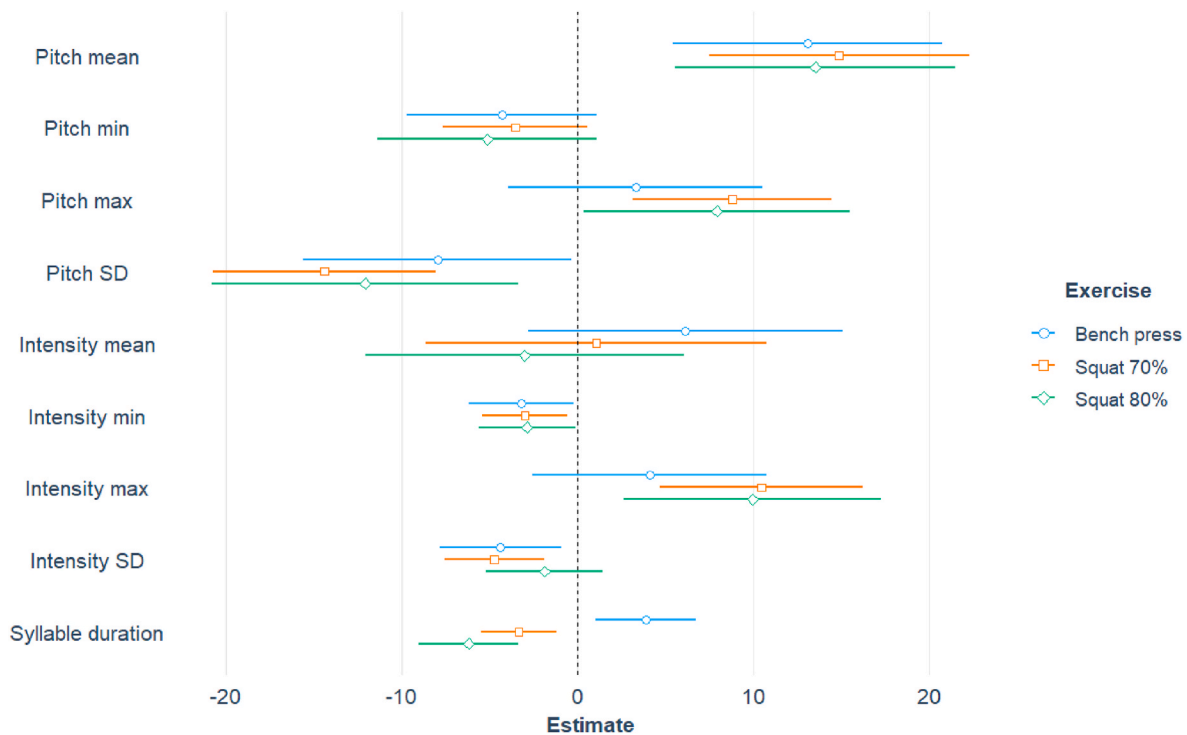


Fig. 4. Mixed model results for the prediction of failure proximity by prosodic features as fixed effects for bench press, squat with 70% 1RM and squat with 80% 1RM exercises (bench press: $n = 19$, $nobs = 710$; squat 70%: $n = 20$, $nobs = 971$; squat 80%: $n = 20$, $nobs = 607$). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration. The circles and squares represent the regression coefficients of the analyses of the first and second subsets, respectively. The error bars represent 95% confidence intervals.

Table 1

Mixed model results for scaled predictors. Prediction of affective valence by speech features as fixed effects, during exercise performance for each of the two subsets ($n = 10$ each; $nobs = 1086$ and $nobs = 1202$, respectively). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration.

Speech feature	<i>b</i>	95% CI	SE	<i>t</i> -value	<i>df</i>	<i>p</i> -value	R^2_{β}
<i>First subset</i>							
Intercept	1.926**	[1.009, 2.829]	0.469	4.111	9.089	0.003	–
Pitch mean (Hz)	–0.926***	[–1.384, –0.497]	0.218	–4.251	19.668	<0.001	0.456
Pitch min (Hz)	0.618***	[0.351, 0.887]	0.132	4.692	1057.652	<0.001	0.02
Pitch max (Hz)	–0.703***	[–0.994, –0.397]	0.157	–4.484	1045.102	<0.001	0.019
Pitch SD (Hz)	1.121***	[0.778, 1.454]	0.178	6.296	922.729	<0.001	0.041
Intensity mean (dB)	–0.375	[–0.957, 0.254]	0.315	–1.189	12.988	0.256	0.091
Intensity min (dB)	0.162*	[–0.014, 0.302]	0.071	2.283	1057.744	0.023	0.005
Intensity max (dB)	–0.589***	[–0.902, –0.264]	0.162	–3.637	1063.568	<0.001	0.012
Intensity SD (dB)	0.412***	[0.265, 0.569]	0.076	5.384	1064.769	<0.001	0.026
Syllable duration (s)	0.574**	[0.243, 0.924]	0.172	3.346	8.779	0.009	0.503
<i>Second subset</i>							
Intercept	2.416**	[1.275, 3.466]	0.577	4.191	8.756	0.002	–
Pitch mean (Hz)	–1.114**	[–1.809, –0.456]	0.356	–3.126	11.981	0.009	0.415
Pitch min (Hz)	0.374**	[0.163, 0.616]	0.113	3.300	1161.664	0.001	0.009
Pitch max (Hz)	–0.917***	[–1.271, –0.603]	0.169	–5.425	1162.982	<0.001	0.025
Pitch SD (Hz)	1.279***	[0.924, 1.661]	0.180	7.118	1148.703	<0.001	0.042
Intensity mean (dB)	0.142	[–0.230, 0.478]	0.180	0.789	36.735	0.435	0.016
Intensity min (dB)	0.152**	[0.037, 0.272]	0.058	2.612	1128.383	0.009	0.006
Intensity max (dB)	–0.473**	[–0.768, –0.147]	0.155	–3.055	1167.949	0.002	0.008
Intensity SD (dB)	0.240**	[0.065, 0.388]	0.079	3.039	1159.771	0.002	0.008
Syllable duration (s)	0.306	[0.040, 0.559]	0.137	2.234	8.730	0.053	0.326

Note: CI = confidence interval, SE = standard error, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

relationship between affective valence and the maximum intensity ($p = 0.005$, $p < 0.001$ and $p = 0.049$ in the bench press, squat 70%, and squat 80% exercises, respectively) (see Fig. 2 and Table 2). Last, the faster the speech rate was in the squat 70% and squat 80% exercises, the higher the affective valence ratings ($p < 0.001$ in both exercises). The same trend was found in the bench-press exercise ($p = 0.063$). This indicates that the affective state is also expressed in the speech rate in different exercises,

at least to some degree (see Fig. 2 and Table 2).

Proximity to task failure. There was an association between proximity to task failure and two of the pitch characteristics (see Fig. 3 and Table 3). In the two subsets, we found a positive relationship between proximity to task failure and pitch mean ($p = 0.037$ and $p < 0.001$ in the first and second subset, respectively) as well as between proximity to task failure and the maximum pitch ($p = 0.005$ and $p = 0.001$ in the first

Table 2

Mixed model results for scaled predictors. Prediction of affective valence by speech features as fixed effects, during each exercise performance (bench-press: $n = 19$, $n_{\text{obs}} = 710$; squat 70%: $n = 20$, $n_{\text{obs}} = 971$; squat 80%: $n = 20$, $n_{\text{obs}} = 607$). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration.

Speech feature	<i>b</i>	95% CI	SE	<i>t</i> -value	<i>df</i>	<i>p</i> -value	R^2_{β}
<i>Bench-press</i>							
Intercept	2.643***	[1.456, 3.474]	0.425	17.91	6.221	<0.001	–
Pitch mean (Hz)	–1.043**	[–1.671, –0.435]	0.342	23.309	–3.048	0.006	0.273
Pitch min (Hz)	0.187	[–0.080, 0.447]	0.133	621.461	1.41	0.159	0.003
Pitch max (Hz)	–0.383*	[–0.844, –0.007]	0.185	679.333	–2.073	0.039	0.006
Pitch SD (Hz)	0.726**	[0.268, 1.177]	0.213	683.682	3.414	0.001	0.017
Intensity mean (dB)	0.108	[–0.523, 0.534]	0.262	28.02	0.412	0.683	0.006
Intensity min (dB)	0.194**	[0.037, 0.340]	0.069	663.718	2.833	0.005	0.012
Intensity max (dB)	–0.423**	[–0.684, –0.059]	0.16	624.414	–2.652	0.008	0.011
Intensity SD (dB)	0.169	[0.015, 0.304]	0.086	591.349	1.966	0.050	0.006
Syllable duration (s)	0.098	[–0.106, 0.313]	0.086	19.635	1.139	0.269	0.059
<i>Squat 70%</i>							
Intercept	2.006***	[1.352, 2.859]	0.436	18.28	4.602	<0.001	–
Pitch mean (Hz)	–1.168***	[–1.646, –0.650]	0.243	43.606	–4.81	<0.001	0.339
Pitch min (Hz)	0.609***	[0.264, 0.852]	0.141	912.264	4.328	<0.001	0.02
Pitch max (Hz)	–0.877***	[–1.249, –0.564]	0.183	858.943	–4.792	<0.001	0.026
Pitch SD (Hz)	1.355***	[0.998, 1.804]	0.202	695.555	6.719	<0.001	0.061
Intensity mean (dB)	–0.039	[–0.502, 0.484]	0.251	38.664	–0.156	0.877	0.001
Intensity min (dB)	0.256**	[0.129, 0.403]	0.074	874.415	3.453	0.001	0.013
Intensity max (dB)	–0.63**	[–1.014, –0.199]	0.188	895.845	–3.355	0.001	0.012
Intensity SD (dB)	0.416***	[0.212, 0.638]	0.093	878.462	4.462	<0.001	0.022
Syllable duration (s)	0.539**	[0.170, 0.821]	0.155	18.421	3.49	0.003	0.378
<i>Squat 80%</i>							
Intercept	1.835	[1.028, 2.738]	0.425	16.295	4.32	0.001	–
Pitch mean (Hz)	–1.03***	[–1.599, –0.414]	0.265	42.714	–3.891	<0.001	0.256
Pitch min (Hz)	0.491**	[0.142, 0.857]	0.161	558.157	3.058	0.002	0.016
Pitch max (Hz)	–0.636**	[–1.129, –0.179]	0.208	560.986	–3.063	0.002	0.016
Pitch SD (Hz)	0.991***	[0.342, 1.500]	0.241	526.974	4.103	<0.001	0.031
Intensity mean (dB)	0.003	[–0.503, 0.547]	0.271	82.251	0.01	0.992	<0.001
Intensity min (dB)	0.071	[–0.121, 0.269]	0.091	572.593	0.789	0.43	0.001
Intensity max (dB)	–0.46*	[–0.866, –0.079]	0.215	550.562	–2.134	0.033	0.008
Intensity SD (dB)	0.159	[–0.058, 0.341]	0.091	567.704	1.739	0.083	0.005
Syllable duration (s)	0.598**	[0.260, 0.972]	0.176	20.463	3.403	0.003	0.345

Note: CI = confidence interval, SE = standard error, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 3

Mixed model results for scaled predictors. Prediction of proximity to task failure by speech features as fixed effects, during exercise performance for each of the two subsets ($n = 10$ each; $n_{\text{obs}} = 1086$ and $n_{\text{obs}} = 1202$, respectively). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration.

Speech feature	<i>b</i>	95% CI	SE	<i>t</i> -value	<i>df</i>	<i>p</i> -value	R^2_{β}
<i>First subset</i>							
Intercept	52.520***	[46.193, 58.586]	3.123	16.817	6.674	<0.001	–
Pitch mean (Hz)	7.914*	[0.877, 15.254]	3.531	2.241	18.009	0.037	0.207
Pitch min (Hz)	–3.871	[–8.525, 0.988]	2.256	–1.716	974.556	0.086	0.003
Pitch max (Hz)	7.448**	[2.531, 12.885]	2.689	2.770	1020.655	0.005	0.007
Pitch SD (Hz)	–9.397**	[–15.466, –3.263]	3.003	–3.130	689.462	0.001	0.014
Intensity mean (dB)	10.651*	[1.880, 18.888]	4.187	2.544	16.289	0.021	0.268
Intensity min (dB)	–4.229***	[–6.749, –1.808]	1.210	–3.496	1026.944	<0.001	0.012
Intensity max (dB)	3.436	[–1.885, 8.990]	2.758	1.246	1000.864	0.212	0.002
Intensity SD (dB)	–5.254***	[–7.908, –2.709]	1.301	–4.039	974.609	<0.001	0.016
Syllable duration (s)	–1.441	[–7.083, 3.908]	2.850	–0.506	8.843	0.625	0.025
<i>Second subset</i>							
Intercept	52.460***	[42.732, 60.919]	4.734	11.082	7.875	<0.001	–
Pitch mean (Hz)	18.499***	[12.279, 24.992]	3.044	6.078	18.995	<0.001	0.627
Pitch min (Hz)	–3.828*	[–7.444, –0.057]	1.789	–2.140	1084.460	0.033	0.004
Pitch max (Hz)	9.139**	[3.920, 14.353]	2.661	3.434	1118.960	0.001	0.01
Pitch SD (Hz)	–16.414***	[–22.250, –11.058]	2.816	–5.828	1021.610	<0.001	0.032
Intensity mean (dB)	–3.733	[–9.913, –2.699]	3.314	–1.127	22.721	0.272	0.051
Intensity min (dB)	–3.819***	[–5.630, –2.115]	0.929	–4.109	1155.250	<0.001	0.014
Intensity max (dB)	5.769*	[0.690, 10.432]	2.467	2.339	1179.060	0.020	0.005
Intensity SD (dB)	–0.109	[–2.570, 2.380]	1.259	–0.087	1180.170	0.931	<0.001
Syllable duration (s)	0.004	[–5.069, 5.049]	2.535	0.002	9.716	0.999	<0.001

Note: CI = confidence interval, SE = standard error, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

and second subset, respectively). Furthermore, we found a negative relationship between proximity to task failure and pitch maximum in the second subset ($p = 0.033$) and the same trend in the first subset ($p = 0.086$), as well as a similar relationship between proximity to task failure

and the pitch SD ($p = 0.001$ and $p < 0.001$ in the first and second subset, respectively). In addition, one intensity characteristic predicted proximity to task failure. We found a negative relationship between proximity to task failure and minimum intensity ($p < 0.001$ in both subsets)

(see Fig. 3 and Table 3). These findings suggest that prosodic features, and mainly pitch characteristics, can consistently predict the level of effort – the proximity to task failure during resistance exercises.

Analyses of each exercise separately, when merging the two subsets, revealed that pitch mean and SD predicted proximity to task failure in all three exercises (see Fig. 4 and Table 4). We found a positive relationship between proximity to task failure and the mean pitch ($p = 0.006$, $p < 0.001$ and $p = 0.001$ in the bench press, squat 70% and squat 80% exercises, respectively). We also found a negative relationship between proximity to task failure and pitch SD ($p = 0.014$, $p < 0.001$ and $p = 0.007$ in the bench press, squat 70%, and squat 80% exercises, respectively), as well as the minimum intensity ($p = 0.035$, $p = 0.016$ and $p = 0.039$ in the bench press, squat 70%, and squat 80% exercises, respectively). These findings suggest that certain prosodic features can predict effort level across different exercises.

4. Discussion

We examined data from a previously conducted experiment (Emanuel et al., 2020) in which participants had to perform three sets of resistance exercises to task failure in three different sessions. After each repetition, participants vocally rated how they felt (i.e., affective valence rating), and their answers were recorded using a tap-on microphone. We extracted the participants' prosodic features from each of the recordings and tested whether they could predict affective valence ratings and proximity to task failure in two separate subsets of the data, and for each exercise separately. This work, to the best of our knowledge, provides the first evidence for the link between prosodic features, affective state and effort level during bouts of exercise.

We found that participants' pitch was deeply manifested in affective

ratings related to effort exertion. In the two subsets of the data and in all the three exercises, the mean and maximum pitch, as well as the SD of the pitch, predicted the affective valence, showing the robustness of these pitch characteristics. Moreover, the results in both subsets and all three exercises also suggest that the louder the maximal intensity of the speech during affective ratings, the worse a person feels.

Notably, pitch characteristics predicted both affective valence and task failure during bouts of exercise across the two datasets. This suggests that pitch characteristics are robustly manifested in physical effort-related tasks. This generalization from the affective domain to the physical domain implies that pitch characteristics may predict effort-related phenomena in a wide range of contexts. Remarkably, the mean pitch was negatively related to affective valence and positively related to task failure in both subsets and across all three exercises and in each exercise separately. In addition, the SD of the pitch showed an opposite relationship where it was positively related to affective valence and negatively related to task failure in both subsets and in each exercise separately. Thus, prosodic features and especially the mean and SD of the pitch enable an indirect estimation of effort and effort-related affective valence in various physical endeavors. Specifically, when the mean and SD of the pitch are higher and lower, respectively, more effort is exerted, and accordingly the affective valence decreases. It is possible that the decrease in the SD of the pitch closer to task failure stems from a ceiling effect – the mean pitch increases during the set closer to the highest possible pitch a person can produce. The distribution of pitch values, therefore, cannot be as wide as it was at the beginning of the set.

The findings regarding the mean pitch align with previous findings showing that pitch is the most perceptually salient acoustic property of the voice (Aung & Puts, 2020). The pitch can convey implied psychological and emotional information (Bailen et al., 2019; Cowen & Keltner,

Table 4

Mixed model of results for scaled predictors. Prediction of proximity to task failure by speech features as fixed effects, during each exercise performance (bench-press: $n = 19$, $n_{\text{obs}} = 710$; squat 70%: $n = 20$, $n_{\text{obs}} = 971$; squat 80%: $n = 20$, $n_{\text{obs}} = 607$). All analyses included a random intercept and random slopes for mean pitch, mean intensity, and syllable duration.

Speech feature	<i>b</i>	95% CI	SE	<i>t</i> -value	<i>df</i>	<i>p</i> -value	R^2_{β}
<i>Bench-press</i>							
Intercept	58.563***	[52.482, 65.801]	3.067	9.607	19.094	<0.001	–
Pitch mean (Hz)	14.538**	[6.740, 22.879]	4.112	34.965	3.535	0.001	0.256
Pitch min (Hz)	–4.778	[–10.091, 1.092]	2.778	584.845	–1.72	0.086	0.005
Pitch max (Hz)	3.666	[–3.934, 11.905]	3.893	674.93	0.942	0.347	0.001
Pitch SD (Hz)	–9.41*	[–20.240, –1.637]	4.228	503.196	–2.226	0.026	0.01
Intensity mean (dB)	7.492	[–0.401, 15.998]	4.637	27.365	1.616	0.118	0.084
Intensity min (dB)	–3.228*	[–6.248, –0.240]	1.472	671.188	–2.194	0.029	0.007
Intensity max (dB)	3.033	[–2.770, 8.313]	3.438	634.25	0.882	0.378	0.001
Intensity SD (dB)	–4.141*	[–7.283, –0.801]	1.837	653.757	–2.254	0.024	0.008
Syllable duration (s)	4.398*	[1.313, 8.988]	1.828	20.673	2.406	0.026	0.209
<i>Squat 70%</i>							
Intercept	49.851	[41.729, 57.758]	4.217	16.549	11.821	<0.001	–
Pitch mean (Hz)	13.504***	[7.160, 19.989]	3.414	34.463	3.956	<0.001	0.303
Pitch min (Hz)	–3.599	[–7.732, 1.052]	2.127	865.619	–1.692	0.091	0.003
Pitch max (Hz)	6.791*	[1.321, 11.319]	2.749	812.554	2.47	0.014	0.007
Pitch SD (Hz)	–11.656***	[–18.153, –5.299]	2.995	493.504	–3.892	<0.001	0.03
Intensity mean (dB)	2.239	[–7.498, 14.482]	4.677	28.785	0.479	0.636	0.008
Intensity min (dB)	–2.831*	[–4.904, –0.954]	1.12	881.772	–2.528	0.012	0.007
Intensity max (dB)	8.282**	[1.363, 16.407]	2.861	934.594	2.895	0.004	0.009
Intensity SD (dB)	–4.43**	[–7.972, –1.809]	1.411	864.456	–3.141	0.002	0.011
Syllable duration (s)	–0.984	[–7.311, –5.141]	2.618	18.941	–0.376	0.711	0.007
<i>Squat 80%</i>							
Intercept	52.084***	[45.485, 59.066]	3.635	11.724	14.329	<0.001	–
Pitch mean (Hz)	11.176**	[6.314, 17.567]	3.256	154.539	3.433	0.001	0.07
Pitch min (Hz)	–5.67*	[–11.329, –0.439]	2.675	555.729	–2.12	0.034	0.008
Pitch max (Hz)	8.143*	[0.183, 15.281]	3.411	555.945	2.387	0.017	0.01
Pitch SD (Hz)	–10.94**	[–20.773, –1.368]	3.906	555.127	–2.801	0.005	0.014
Intensity mean (dB)	–3.798	[–14.283, 7.370]	4.949	51.17	–0.767	0.446	0.011
Intensity min (dB)	–2.85	[–6.342, 0.690]	1.495	574.821	–1.906	0.057	0.006
Intensity max (dB)	10.567**	[3.962, 17.874]	3.62	561.933	2.919	0.004	0.015
Intensity SD (dB)	–2.458	[–5.795, 0.361]	1.522	560.368	–1.614	0.107	0.005
Syllable duration (s)	–5.453	[–12.779, 0.271]	3.08	19.378	–1.77	0.092	0.132

Note: CI = confidence interval, SE = standard error, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2017; Dietrich et al., 2019; Morningstar et al., 2017; Stern et al., 2021) and is crucial for the interpretation of emotional intent (Morningstar et al., 2017; Pell et al., 2009), such as the expression of anger, fear, happiness, sadness and disgust (Cao et al., 2014). Interestingly, during most of the aforementioned emotions, the vocal emotional expression is associated with increased mean pitch. This may be explained by sympathetic arousal, such as when experiencing a wide array of emotions and during effort exertion, which leads to an increase in the mean pitch during vocal expression (Banse & Scherer, 1996; Rodero, 2022).

We would like to emphasize that during physical effort, as in all the three tasks in this study (lifting weights), the more effort invested during bouts of exercise, the higher the cardiovascular strain. Thus, the degree of breathlessness expresses the effort level one experiences by influencing the prosody and, more specifically, the pitch. The production of speech requires energy, and physical effort may be reflected in the pitch as the energy in the exhalation influences the pitch (Primov-Fever et al., 2014). Pitch also plays a role in the perceptions of characteristics related to social power, such as dominance and leadership (Aung & Puts, 2020). Higher pitch reflects low competence (Klofstad et al., 2012) and greater stressfulness (Quinto et al., 2013). This may imply that an elevation in pitch during effort exertion reflects stress – the attempt to keep performing a task though in the face of hardship. In addition, this hard-to-prevent expression of physical effort may be an important social signal that may help others to be mentally and physically prepared for an effortful situation. It is possible that similar to the high pitch when crying (Hepburn, 2004; Hepburn & Potter, 2007), high pitch during effort exertion may serve as a cue of distress to the environment, which may aid people to support their recovery later on.

Our findings suggest that by recording people speaking, it may be possible to overcome certain inaccuracies stemming from individual variability in reporting. For example, certain people may find it hard to admit difficulty more than others, which may lead to unreliable self-reports regarding effort-exertion. More broadly, any deficit in the ability to articulate oneself may lead to inaccurate effort estimation. This may be even more prominent in a case of relevant psychopathology such as autism (Chan et al., 2005).

Predicting effort levels by prosodic features enables an ecological measurement of effort in a wide array of contexts of everyday settings and expands the scope in which effort can be examined. Pitch characteristics and specifically the mean pitch and the SD of the pitch can be used both in the field and in research contexts in order to estimate the effort level of participants during a strenuous task (Hackett et al., 2012, 2017; Zourdos et al., 2016). For example, people working out at the gym may estimate how close they are to their maximal capacity simply by recording themselves saying a word via a designated app. More broadly, affective responses to exercises were found to predict adherence to training programs in untrained populations (Lee et al., 2016; Williams et al., 2008, 2016), and were suggested as the main consideration in exercise prescription (Ekkekakis, 2009; Ekkekakis et al., 2000, 2011; Greene & Petruzzello, 2015; Hutchinson et al., 2020). Our findings suggest that trainees and coaches can monitor affective states during exercise through speech recordings and adjust exercise intensity in a manner that may increase future adherence to exercise programs or prevent burnout in competitive athletes. Specifically, coaches can directly assess trainees' affective state during each exercise via their prosodic features, enabling them to cease exercise performance before the affective state drops.

Beyond the relevance to everyday settings, prosodic features may allow further elaboration of the predictions of effort allocation models (Emanuel, 2019; Inzlicht et al., 2014; Kurzban et al., 2013; Shenhav et al., 2017). For example, certain models predict a decline in the affective state over time (e.g., Kurzban et al., 2013), while others suggest that the affective state may also increase as people approach their deadline (e.g., Emanuel, Katzir, & Liberman, 2022). By employing prosodic features to assess the affective states and the level of effort during prolonged effort exertion, it would be possible to avoid asking

participants to fill questionnaires. This can further provide support for one account or the other with less methodological difficulties.

Another issue is the ongoing debate regarding the underpinnings of the perception of effort. According to one approach, the perception of effort is calculated from a neural copy of a motor command signal (the corollary discharge model; Marcora & Staiano, 2010; Morree et al., 2012). A competing model, in contrast, claims that the perception of effort is a multi-factored neuro-computational mechanism, which its only function is to maintain homeostasis (the central governor model; Noakes, 2012). The theoretical inconsistencies may stem at least partially from the context-dependent measures such as the specific questionnaires that are used to measure the perception of effort (Emanuel et al., 2020). We suggest that using prosodic features instead, or along the perception of effort questionnaires, might better capture effort during exercise and its related affective response. Being more objective than self-report questionnaires, prosodic features may aid in elucidating which model of effort perception is the most accurate.

Further research is needed to determine the generalizability our results to the non-trained population and the generalizability of prosodic features for other physical effort predictions as well as non-physical predictions. For example, prosodic features may predict rating of perceived exertion and arousal. An important potential question would be whether prosodic features and especially the mean pitch and the SD of the pitch can predict effort level in tasks of a more cognitive nature. For example, testing whether vocal features reflect the level of cognitive effort through different working-memory loads.

To conclude, we found that prosodic features predicted both effort-related affective valence ratings and proximity to task failure across two subsets of experimental data. Specifically, the mean pitch and the SD of the pitch reflected the effort-related affective valence ratings and proximity to task failure in both subsets and when examining each of the three physical exercises separately. This suggests that prosodic features, and especially the mean pitch and the SD of the pitch, can be used to monitor effort-level and related affect across a variety of physical exercises and help uncover the nature of physical effort in its different manifestations.

Author contributions

IR and AE conceived the idea, IR extracted the speech features, AE cleaned and formatted the data, AE and IR analyzed the data and wrote the paper.

Open practices statement

Data and R syntax for the analyses are available at <https://bit.ly/3jlJtGN>.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have shared the data via a link to an OSF server within the manuscript

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.mhpa.2023.100559>.

References

- Adams, S. A. (2005). The effect of social desirability and social approval on self-reports of physical activity. *American Journal of Epidemiology*, 161(4), 389–398. <https://doi.org/10.1093/aje/kwi054>
- Anikin, A., & Lima, C. F. (2017). Perceptual and acoustic differences between authentic and acted nonverbal emotional vocalizations. *The Quarterly Journal of Experimental Psychology*, 1–21. <https://doi.org/10.1080/17470218.2016.127097>
- Anikin, A., & Persson, T. (2017). Nonlinguistic vocalizations from online amateur videos for emotion research: A validated corpus. *Behavior Research Methods*, 49(2), 758–771. <https://doi.org/10.3758/s13428-016-0736-y>
- Anikin, A., & Reby, D. (2022). Ingressive phonation conveys arousal in human nonverbal vocalizations. *Bioacoustics*, 1(6), 680–695. <https://doi.org/10.1080/09524622.2022.2039295>
- Aung, T., & Puts, D. (2020). Voice pitch: A window into the communication of social power. *Current Opinion in Psychology*, 33, 154–161. <https://doi.org/10.1016/j.copsyc.2019.07.028>
- Bailen, N. H., Green, L. M., & Thompson, R. J. (2019). Understanding emotion in adolescents: A review of emotional frequency, intensity, instability, and clarity. *Emotion Review*, 11(1), 63–73. <https://doi.org/10.1177/1754073918768878>
- Banase, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*, 70(3), 614–636. <https://doi.org/10.1037/0022-3514.70.3.614>
- Bliese, P. D., & Ployhart, R. E. (2002). Growth modeling using random coefficient models: Model building, testing, and illustrations. *Organizational Research Methods*, 5(4), 362–387. <https://doi.org/10.1177/109442802237116>
- Boersma, P., & Weenink, D. (2021). Praat: Doing phonetics by computer [Computer program] (Version 6.1. 24). www.praat.org. Retrieved from.
- Brenner, P. S., & DeLamater, J. (2016). Lies, damned lies, and survey self-reports? Identity as a cause of measurement bias. *Social Psychology Quarterly*, 79(4), 333–354. <https://doi.org/10.1177/0190272516628298>
- Cao, H., Benus, S., Gur, R. C., Verma, R., & Nenkova, A. (2014). Prosodic cues for emotion: Analysis with discrete characterization of intonation. *Speech Prosody (Urbana, Ill.)*, 2014, 130–134. <https://doi.org/10.21437/SpeechProsody.2014-14>
- Chan, A. S., Cheung, J., Leung, W. W. M., Cheung, R., & Cheung, M. (2005). Verbal expression and comprehension deficits in young children with autism. *Focus on Autism and Other Developmental Disabilities*, 20(2), 117–124. <https://doi.org/10.1177/10883576050200020201>
- Chen, J., & Wu, C. (2022). On the relationship between well-being and exercise adherence for children and adolescents: A systematic mini review. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.900287>
- Cowen, A. S., & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, 114(38), E7900–E7909. <https://doi.org/10.1073/pnas.1702247114>
- De Jong, B. A., & Elfring, T. (2010). How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort. *Academy of Management Journal*, 53(3), 535–549. <https://doi.org/10.5465/amj.2010.51468649>
- De Waele, A., Claeys, A. S., & Cauterberg, V. (2019). The organizational voice: The importance of voice pitch and speech rate in organizational crisis communication. *Communication Research*, 46(7), 1026–1049. <https://doi.org/10.1177/0093650217692911>
- Dickinson, D., & Villeval, M.-C. (2008). Does monitoring decrease work effort?: The complementarity between agency and crowding-out theories. *Games and Economic Behavior*, 63(1), 56–76. <https://doi.org/10.1016/j.geb.2007.08.004>
- Dietrich, B. J., Hayes, M., & O'Brien, D. Z. (2019). Pitch perfect: Vocal pitch and the emotional intensity of congressional speech. *American Political Science Review*, 113(4), 941–962. <https://doi.org/10.1017/S0003055419000467>
- Ekkekakis, P. (2009). Let them roam free? Physiological and psychological evidence for the potential of self-selected exercise intensity in public health. *Sports Medicine*, 39(10), 857–888. <https://doi.org/10.2165/11315210-000000000-00000>
- Ekkekakis, P., Hall, E. E., VanLanduyt, L. M., & Petruzzello, S. J. (2000). Walking in (affective) circles: Can short walks enhance affect? *Journal of Behavioral Medicine*, 23(3), 245–275. <https://doi.org/10.1023/A:1005558025163>
- Ekkekakis, P., Parfitt, G., & Petruzzello, S. J. (2011). The pleasure and displeasure people feel when they exercise at different intensities: Decennial update and progress towards a tripartite rationale for exercise intensity prescription. *Sports Medicine*, 41(8), 641–671. <https://doi.org/10.2165/11590680-000000000-00000>
- Emanuel, A. (2019). Perceived impact as the underpinning mechanism of the end-sprint and u-shape pacing patterns. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01082>
- Emanuel, A., Katzir, M., & Liberman, N. (2022). Why do people increase effort near a deadline? An opportunity-cost model of goal gradients. *Journal of Experimental Psychology: General*.
- Emanuel, A., Smukas, I. R., & Halperin, I. (2020). How one feels during resistance exercises: A repetition-by-repetition analysis across exercises and loads. *International Journal of Sports Physiology and Performance*, 16(1), 135–144. <https://doi.org/10.1123/ijsp.2019-0733>
- Fafchamps, M., & Labonne, J. (2017). Using split samples to improve inference on causal effects. *Political Analysis*, 25(4), 465–482. <https://doi.org/10.1017/pan.2017.22>
- Filippa, M., Lima, D., Grandjean, A., Labbé, C., Coll, S. Y., Gentaz, E., & Grandjean, D. M. (2022). Emotional prosody recognition enhances and progressively complexifies from childhood to adolescence. *Scientific Reports*, 12, Article 17144. <https://doi.org/10.1038/s41598-022-21554-0>
- Gharavian, D., Sheikhan, M., & Janipour, M. (2010). Pitch in emotional speech and emotional speech recognition using pitch frequency. *Majlesi Journal of Electrical Engineering*, 4(1), 19.
- Greene, D. R., & Petruzzello, S. J. (2015). More isn't necessarily better: Examining the intensity-affect-enjoyment relationship in the context of resistance exercise. *Sport, Exercise, and Performance Psychology*, 4(2), 75–87. <https://doi.org/10.1037/spy0000030>
- Guidi, A., Gentili, C., Scilingo, E. P., & Vanello, N. (2019). Analysis of speech features and personality traits. *Biomedical Signal Processing and Control*, 51, 1–7. <https://doi.org/10.1016/j.bspc.2019.01.027>
- Hackett, D. A., Cobley, S. P., Davies, T. B., Michael, S. W., & Halaki, M. (2017). Accuracy in estimating repetitions to failure during resistance exercise. *The Journal of Strength & Conditioning Research*, 31(8), 2162–2168. <https://doi.org/10.1519/JSC.0000000000001683>
- Hackett, D. A., Johnson, N. A., Halaki, M., & Chow, C.-M. (2012). A novel scale to assess resistance-exercise effort. *Journal of Sports Sciences*, 30(13), 1405–1413. <https://doi.org/10.1080/02640414.2012.710757>
- Haines, M., Broom, D., Gillibrand, W., & Stephenson, J. (2020). Effects of three low-volume, high-intensity exercise conditions on affective valence. *Journal of Sports Sciences*, 38(2), 121–129. <https://doi.org/10.1080/02640414.2019.1684779>
- Hardy, C. J., & Rejeski, W. J. (1989). Not what, but how one feels: The measurement of affect during exercise. *Journal of Sport & Exercise Psychology*, 11(3), 304–317. <https://doi.org/10.1123/jsep.11.3.304>
- Henton, C. (1995). Pitch dynamism in female and male speech. *Language & Communication*, 15(1), 43–61. [https://doi.org/10.1016/0271-5309\(94\)00011-Z](https://doi.org/10.1016/0271-5309(94)00011-Z)
- Hepburn, A. (2004). Crying: Notes on description, transcription, and interaction. *Research on Language & Social Interaction*, 37(3), 251–290. <https://doi.org/10.1207/s15327973rlsi3703.1>
- Hepburn, A., & Potter, J. (2007). Crying receipts: Time, empathy, and institutional practice. *Research on Language and Social Interaction*, 40(1), 89–116. <https://doi.org/10.1080/08351810701331299>
- Hutchinson, J. C., Zenko, Z., Santich, S., & Dalton, P. C. (2020). Increasing the pleasure and enjoyment of exercise: A novel resistance-training protocol. *Journal of Sport & Exercise Psychology*, 42(2), 143–152. <https://doi.org/10.1123/jsep.2019-0089>
- Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, 18(3), 127–133. <https://doi.org/10.1016/j.tics.2013.12.009>
- Jaeger, B. C., Edwards, L. J., Das, K., & Sen, P. K. (2017). An R^2 statistic for fixed effects in the generalized linear mixed model. *Journal of Applied Statistics*, 44(6), 1086–1105. <https://doi.org/10.1080/02664763.2016.1193725>
- de Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, 41(2), 385–390. <https://doi.org/10.3758/BRM.41.2.385>
- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698–2704. <https://doi.org/10.1098/rspb.2012.0311>
- Koolagudi, S. G., & Krothapalli, R. S. (2011). Two stage emotion recognition based on speaking rate. *International Journal of Speech Technology*, 14(1), 35–48. <https://doi.org/10.1007/s10772-010-9085-x>
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, 36(6), 661–679. <https://doi.org/10.1017/S0140525X12003196>
- Lee, H. H., Emerson, J. A., & Williams, D. M. (2016). The exercise-affect-adherence pathway: An evolutionary perspective. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2016.01285>, 07.
- Lemyre, P.-N., Treasure, D. C., & Roberts, G. C. (2006). Influence of variability in motivation and affect on elite athlete burnout susceptibility. *Journal of Sport & Exercise Psychology*, 28(1), 32–48. <https://doi.org/10.1123/jsep.28.1.32>
- Mandolesi, L., Polverino, A., Montuori, S., Foti, F., Ferraioli, G., Sorrentino, P., & Sorrentino, G. (2018). Effects of physical exercise on cognitive functioning and wellbeing: Biological and psychological benefits. *Frontiers in Psychology*, 9. <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00509>
- Manson, J. H., Bryant, G. A., Gervais, M. M., & Kline, M. A. (2013). Convergence of speech rate in conversation predicts cooperation. *Evolution and Human Behavior*, 34(6), 419–426. <https://doi.org/10.1016/j.evolhumbehav.2013.08.001>
- Marcora, S. M., & Staiano, W. (2010). The limit to exercise tolerance in humans: Mind over muscle? *European Journal of Applied Physiology*, 109(4), 763–770. <https://doi.org/10.1007/s00421-010-1418-6>
- Micklewright, D., St Clair Gibson, A., Gladwell, V., & Al Salman, A. (2017). Development and Validity of the rating-of-fatigue scale. *Sports Medicine*, 47(11), 2375–2393. <https://doi.org/10.1007/s40279-017-0711-5>
- Morningstar, M., Dirks, M. A., & Huang, S. (2017). Vocal cues underlying youth and adult portrayals of socio-emotional expressions. *Journal of Nonverbal Behavior*, 41(2), 155–183. <https://doi.org/10.1007/s10919-017-0250-7>
- Morree, H. M., Klein, C., & Marcora, S. M. (2012). Perception of effort reflects central motor command during movement execution: Neurophysiology of perceived effort. *Psychophysiology*, 49(9), 1242–1253. <https://doi.org/10.1111/j.1469-8986.2012.01399.x>
- Noakes, T. D. (2012). The Central Governor Model in 2012: Eight new papers deepen our understanding of the regulation of human exercise performance. *British Journal of Sports Medicine*, 46(1), 1–3. <https://doi.org/10.1136/bjsports-2011-090811>
- Olsen Wayne, O. (1998). Average speech levels and spectra in various speaking/listening conditions. *American Journal of Audiology*, 7(2), 21–25. [https://doi.org/10.1044/1059-0889\(1998\)012](https://doi.org/10.1044/1059-0889(1998)012)
- Pell, M. D., Paulmann, S., Dara, C., Alaseri, A., & Kotz, S. A. (2009). Factors in the recognition of vocally expressed emotions: A comparison of four languages. *Journal of Phonetics*, 37(4), 417–435. <https://doi.org/10.1016/j.wocn.2009.07.005>

- Primov-Fever, A., Lidor, R., Meckel, Y., & Amir, O. (2014). The effect of physical effort on voice characteristics. *Folia Phoniatrica et Logopaedica*, 65(6), 288–293. <https://doi.org/10.1159/000361047>.
- Quinto, L., Thompson, W. F., & Keating, F. L. (2013). Emotional communication in speech and music: The role of melodic and rhythmic contrasts. *Frontiers in Psychology*, 4, 184. <https://doi.org/10.3389/fpsyg.2013.00184>
- Re, D. E., O'Connor, J. J. M., Bennett, P. J., & Feinberg, D. R. (2012). Preferences for very low and very high voice pitch in humans. *PLoS One*, 7(3), Article e32719. <https://doi.org/10.1371/journal.pone.0032719>
- Rodero, E. (2022). Effectiveness, attractiveness, and emotional response to voice pitch and hand gestures in public speaking. *Frontiers in Communication*, 7. <https://doi.org/10.3389/fcomm.2022.869084>
- Sander, D. E., & Scherer, K. R. (2009). *The Oxford companion to emotion and the affective sciences*. Oxford University Press.
- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, 40, 99–124. <https://doi.org/10.1146/annurev-neuro-072116-031526>
- Steele, J., Fisher, J., McKinnon, S., & McKinnon, P. (2016). Differentiation between perceived effort and discomfort during resistance training in older adults: Reliability of trainee ratings of effort and discomfort, and reliability and validity of trainer ratings of trainee effort. *Journal of Trainology*, 6(1), 1–8. <https://doi.org/10.17338/trainology.6.1.1>
- Steele, J., Pinto, M. D., Nosaka, K., & Nuzzo, J. L. (2023). Perceptions of capacity, fatigue, and their psychophysics: Examining construct equivalence and the relationship between actual capacity and perception of capacity during resisted elbow flexion tasks. *Psicológica*, 44(2). <https://doi.org/10.20350/digitalCSIC/15498>
- 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, Article 104092. <https://doi.org/10.1016/j.jrp.2021.104092>
- Symons, A. E., El-Deredy, W., Schwartz, M., & Kotz, S. A. (2016). The functional role of neural oscillations in non-verbal emotional communication. *Frontiers in Human Neuroscience*, 10, 239. <https://doi.org/10.3389/fnhum.2016.00239>
- Williams, D. M., Dunsiger, S., Ciccolo, J. T., Lewis, B. A., Albrecht, A. E., & Marcus, B. H. (2008). Acute affective response to a moderate-intensity exercise stimulus predicts physical activity participation 6 and 12 months later. *Psychology of Sport and Exercise*, 9(3), 231–245. <https://doi.org/10.1016/j.psychsport.2007.04.002>
- Williams, D. M., Dunsiger, S., Emerson, J. A., Gwaltney, C. J., Monti, P. M., & Miranda, R. (2016). Self-paced exercise, affective response, and exercise adherence: A preliminary investigation using ecological momentary assessment. *Journal of Sport & Exercise Psychology*, 38(3), 282–291. <https://doi.org/10.1123/jsep.2015-0232>
- Yokomizo, H., Haccou, P., & Iwasa, Y. (2004). Multiple-year optimization of conservation effort and monitoring effort for a fluctuating population. *Journal of Theoretical Biology*, 230(2), 157–171. <https://doi.org/10.1016/j.jtbi.2004.04.036>
- Zourdos, M. C., Klemp, A., Dolan, C., Quiles, J. M., Schau, K. A., Jo, E., Helms, E., Esgro, B., Duncan, S., Garcia Merino, S., & Blanco, R. (2016). Novel resistance training-specific rating of perceived exertion scale measuring repetitions in reserve. *The Journal of Strength & Conditioning Research*, 30(1), 267. <https://doi.org/10.1519/JSC.0000000000001049>