DOI: 10.61782/fa.2023.0991

forum acusticum 2023

# THE OVERALL PERCEIVED QUALITY RATINGS OF VACUUM CLEANERS IN STATIONARY AND SINGLE-WORKING-CYCLE CONDITIONS

**Toros Senan**[1*] **Stephy Annie Curie**[1] **Willemijn Anna Maria Mattheij**[1]

[1] Sorama BV, Achtseweg Zuid 153H 5651 GW Eindhoven, The Netherlands

## ABSTRACT

Product sound quality has been at the forefront of acoustics, and vacuum cleaner sounds, mostly recorded in stationary working conditions, have been commonly studied in the literature. However, the beginning and end of the vacuum cleaner's working cycle generate non-stationary, spectro-temporally modulated sound, which differs from the sound emitted in stationary working conditions. In the present manuscript, we report a listening experiment using a set of vacuum cleaner sounds recorded in stationary and while in its single-working-cycle conditions: Recording started from the moment when the vacuum cleaner is turned on, continued while it is in stationary mode for 10 seconds, and ended when the emitted sound has fully decayed after turning the device off. The stationary and single-working-cycle conditions were evaluated in a listening experiment concerning their overall perceived quality. The overall perceived quality ratings for both conditions were analyzed with the psychoacoustic parameters. The results showed that while participants' ratings for efficiency and functionality attributes correlated significantly with tonality and sharpness values in stationary mode, the same attributes' ratings correlated significantly with roughness and fluctuation strength values in single-working-cycle condition.

**Keywords:** *sound quality*, *psychoacoustics*, *sound design*, *vacuum cleaners*, *consumer product sound quality*

****Corresponding author***: *toros.senan@sorama.eu.*



## 1. INTRODUCTION

Among household appliances, vacuum cleaners are one of the noisiest products and form one of the most powerful noise sources in every household. Therefore, product manufacturers focused on reducing noise levels produced by the suction units. However, they reached a point where further noise level reduction is hard to achieve or at least not financially feasible [1]. On the other hand, researchers investigated the perceptual aspects of vacuum cleaner sounds and showed that the sound power levels of the products are not sufficient to characterize the users' perception of the product sound (e.g., [2–4]). Furthermore, while two different products might generate similar noise levels, perceived disturbances might be different since investigations on the perception of vacuum cleaner sounds showed that loudness, sharpness, and roughness play an essential role in the annoyance and quality judgments (e.g., [5, 6]). Therefore, reducing the perceived disturbance from the product is possible while keeping the noise levels the same.

Determining the perceived disturbance, annoyance, or overall quality of product sounds, is done via listening tests using perceptual attributes extracted via sensory profiling (for a review, see [7]). Listeners then evaluate these perceptual attributes. Standard methods are the semantic differential technique [8], and multidimensional scaling [9]. Typically, a 7-point Likert scale is used for grading each perceptual attribute [9]. The results of the listening tests are evaluated together with acoustic metrics. In the last two decades, psychoacoustic metrics [10] were included in the analysis stage due to the standardization of some of the metrics (e.g., [11, 12]) and commercial software packages providing tools for psychoacoustic analysis. Researchers proposed models for determining and predicting the products' sound quality using psychoacoustic metrics such as loudness, roughness, sharpness, fluctuation strength, and tonality (e.g., [2, 3, 5]). The combination of listening tests with psychoacoustic analysis was used in different application fields, such as the automotive industry [13], kitchen appliance industry [14], environmental noise studies [15], and vacuum cleaner industry [1–3, 6, 16].

Researchers investigating the acoustical and psychoacoustical indicators of perceptual attributes studied vacuum cleaner sounds in great detail. The effect of measurement setup was investigated by comparing binaural and monaural measurements [5]; new models for suction unit design were evaluated with listening tests, as well as psychoacoustic metrics [17, 18]; quality ratings were statistically predicted from perceptual attribute scores using loudness-matched vacuum cleaners [6]. Most of the known studies focused on the stationary operation mode of vacuum cleaners (e.g., [2, 6]), where a stationary noise is emitted. Very few used the single-working-cycle (SWC) sound as a stimulus in the listening tests [19]. In the present paper, we follow up on the methodologies of the existing literature and introduce a comparison between the stationary and SWC operation modes of loudness-matched, single-microphone recordings of seven bagged canister-type vacuum cleaners using the semantic differential technique. We aim to understand which segments of the sound stimuli play a role in participants' decisions while rating statistically significant perceptual attributes. We use psychoacoustic metrics to guide us through our understanding of vacuum cleaners' acoustical and underlying mechanical components that impact participants' perceptions.

The manuscript is structured as follows: Sec. 2 describes the experimental setup, including recording and experimental procedure; Sec. 3 presents the results and psychoacoustic analysis; Sec. 4 discusses the experimental results and Sec. 5 finalizes the manuscript with concluding remarks.

## 2. EXPERIMENT

### 2.1 Participants

Twenty-one subjects (age range: 25-48) participated in the listening experiment performed at a sound-isolating booth at the Sorama office. Participants reported normal hearing and normal vision during the intake. All participants were volunteering Sorama employees.

### 2.2 Stimuli

Sound recordings of seven vacuum cleaners were recorded in the semi-anechoic chamber at the Sorama acoustic laboratory. Each vacuum cleaner (VC) was placed on the floor, and the CAM64 (can be reached at `https://sorama.eu/products/cam64`) was placed at 165 cm height, angled towards the VC on the floor at a distance of 150 cm from the VC. The hose was attached and positioned along one side of the chamber. The VCs operated at maximum power. Sound recordings were conducted with Sorama Portal (can be found at `https://sorama.eu/products/sorama-portal`).

The recording procedure began with one of the responsible researchers starting the recording outside the chamber and the second researcher turning on the VC by pressing the on-off button. After approximately 10 seconds, the researcher in the chamber pressed the on-off button again, and the recording stopped after the VC sound completely decayed.

Three recordings were conducted per VC, and the one with the stationary section closest to 10 seconds long was chosen for each VC. The chosen VC sound was cut between three and a half and nine and a half seconds. A linear ramp-shaped fade-in and fade-out of 50 ms was applied to the cut section's beginning and end. The resulting segment of the recording was labeled as the *stationary* mode, which contains the stationary section of the emitted sound. For each VC, two sound stimuli were prepared for the experiment: A 6-second long stationary sound cut from the recording and a 20-25 second-long non-stationary sound representing an SWC of the VC (see Fig. 1). In total, seven stationary and seven SWC mode stimuli were prepared for one experimental session. All sounds were loudness matched using N5, calculated using Zwicker's time-varying loudness method [11], where the loudness difference between the stimuli was a maximum of 0.5 Sones, centered around 28 sones. The sound pressure levels corresponding to the matched loudness range between 65-75 dB (A).

### 2.3 Apparatus

The experiment was run on a Dell laptop using Python (3.9). All sound stimuli were prepared in MATLAB at a 44.1-kHz sampling rate to a resolution of 16 bits and played out diotically using a PC soundcard (Steinberg UR22C). The participants were placed in a sound-

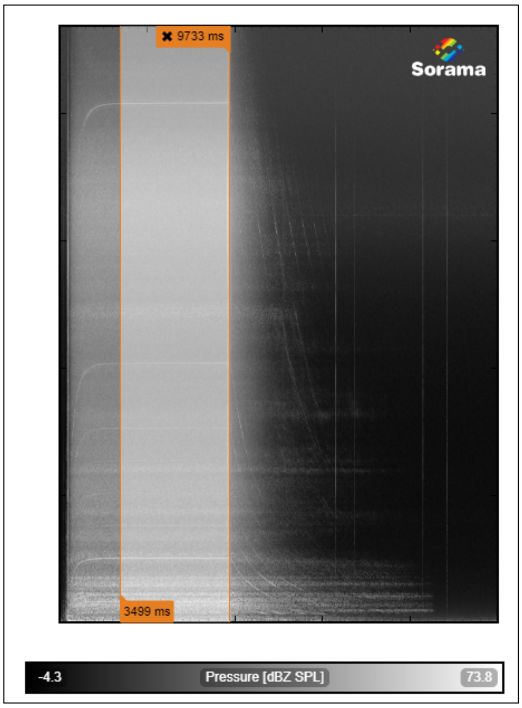


**Figure 1**. Spectrogram showing the recording of one of the VCs. The time selection denoted in the spectrogram represents the stationary mode stimulus of the same VC cut from the SWC recording.

isolating booth at Sorama, and Beyer-Dynamic DT 770 Pro (250 Ohm) headphones were used for playback. The average sound level of the stimuli was calibrated to 65 - 75 dB(A) SPL (see previous section). A computer screen was connected to the laptop.

### 2.4 Procedure

A single trial began after participants clicked the GUI's start button. In each trial, on the top of the page, there was a play button that participants clicked on to listen to the stimulus. Below were 6-word pairs, presented as a 7-point Likert scale, and a 10-point overall perceived quality (OPQ) scale was presented at the bottom of the screen. Participants had to rate each word pair and the OPQ scale to go to the subsequent trial. There were no time constraints for the trial, and they could listen to the sound as often as they wanted. After all of the word pairs and the OPQ scale were rated, a "Next" button appeared on the bottom right side of the GUI. Participants did not have the option to navigate to the previous trial after clicking on the next button.

The semantic differentials are the ones that corresponded to the highest proportion of variance among 15 word pairs in the study of Hülsmeier and colleagues [6]: weak-powerful, inefficient-efficient, broken-functional, rough-smooth, annoying-pleasant, and not howling-howling.

Seven stationary and seven SWC sounds were presented in two different blocks to all participants. The presentation order of blocks and the stimuli within each block were randomized. Each block began with seven training trials, where each VC sound was presented once before the experimental block began. Each stimulus was presented two times during the experimental block. One experimental session took approximately 35-40 mins.

## 3. RESULTS (N = 21)

The perceived quality ratings for each VC in both modes are presented in Fig. 2. It can be observed that participants rated the stationary stimuli differently than spectro-temporally complex SWC mode stimuli. While VC G was rated with the highest quality rating when listened to in stationary mode (M = 65.95 %, SD = 18.35 %), VC C was rated highest by participants when presented in the SWC mode (M = 69.76 %, SD = 15.69 %). Similarly, the lowest ratings were given to VCs D (M = 46.66 %, SD = 22.38 %) and F (M = 48.09 %, SD = 19.03 %) in stationary and SWC modes, respectively.

The two-way repeated measures ANOVA showed that the effect of sound on the OPQ rating was highly significant both in stationary mode ($F(6,120) = 6.08$; $p < 0.05$) and in SWC mode ($F(6,120) = 6.9742$; $p < 0.05$). The OPQ ratings for the same VCs in two conditions were compared by paired t-tests using the Bonferroni correction for seven pairs. The t-tests revealed that the ratings given for VC D (M = 46.68 % SD = 22.38 % in stationary, M = 55.95 %, SD = 20.01 % in SWC mode) and VC F (M = 54.28 % SD = 17.96 % in stationary, M = 48.09 %, SD = 19.03 % in SWC mode) differed significantly between stationary and SWC conditions ($p < 0.05$). We will focus on the sounds of these two VCs throughout the rest of the manuscript to investigate the perceptual differences between the two conditions.

To understand which differences between two conditions played a role in participants' decisions on quality, perceptual attributes were analyzed. The procedure for the analysis began with analyzing the impact of sound

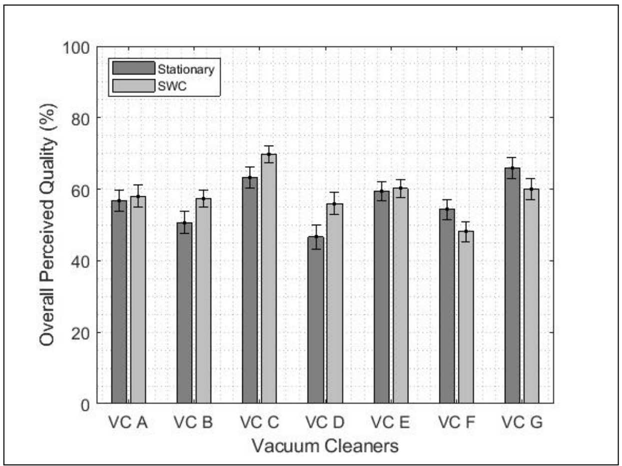


**Figure 2**. The average OPQ ratings are presented as percentages for both operation modes of seven VCs. The error bars show the SEM.

on participants' ratings by two-way repeated measures ANOVA for each perceptual attribute, and the perceptual attributes that did not yield significant interaction ($p > 0.05$) were discarded. For stationary mode, all the perceptual attributes yielded a significant effect of sound on participants' ratings: weak-powerful ($F(6,120) = 2.36$; $p < 0.05$), inefficient-efficient ($F(6,120) = 5.47$; $p < 0.001$) broken-functional ($F(6,120) = 8.30$; $p < 0.001$), rough-smooth ($F(6,120) = 6.23$; $p < 0.001$), annoying-pleasant ($F(6,120) = 32.40$; $p < 0.001$), and not howling-howling ($F(6,120) = 7.95$; $p < 0.001$). For single working cycle mode, all perceptual attributes except the weak-powerful attribute revealed a significant effect of sound on participants' ratings: inefficient-efficient ($F(6,120) = 3.83$; $p < 0.01$), broken-functional ($F(6,120) = 5.58$; $p < 0.001$), rough-smooth ($F(6,120) = 7.52$; $p < 0.001$), annoying-pleasant ($F(6,120) = 22.30$; $p < 0.001$) and not howling-howling ($F(6,120) = 3.83$; $p < 0.01$).

Second, stationary and SWC stimuli of the VCs D and F were compared with paired t-tests for each statistically significant perceptual attribute. The t-tests revealed significant differences for the ratings of weak-powerful ($p < 0.01$), broken-functional ($p < 0.001$), annoying-pleasant ($p < 0.05$), and not howling-howling ($p < 0.001$) attributes given for VC D in stationary and SWC conditions. For VC F, t-tests revealed significant differences for the ratings of inefficient-efficient ($p < 0.05$) and not howling-howling ($p < 0.05$) attributes. Additionally, Pearson correlations between the selected perceptual attributes were calculated for relating each attribute's rating to the OPQ ratings and are given in Tab. 1.

Third, the differences between the stationary and SWC mode stimuli were determined: The segment until the sound reached the stationary status, which contained the sound generated by the on/off switch and the spectro-temporal change until the stationary section, is labeled as the *ascent* section; The switch sound at the end followed by the decay due after the VCs were turned off is labeled as the *descent* section. Psychoacoustic metrics of fluctuation strength (FS), tonality (T), loudness (N), sharpness (S), and roughness (R) were calculated for the stationary mode stimuli and each section of the SWC condition. T and R values are based on the models published in ECMA-418 [12]. FS values are based on the FS model developed for Sorama Portal. N values are based on Zwicker's time-varying loudness model [11], and S values are calculated using Zwicker's time-varying specific loudness values.

The statistical descriptors, mean, RMS, max, percentile 90, and percentile 95, were calculated for each psychoacoustic metric. The Pearson correlation between each statistical descriptive of every psychoacoustic metric for stationary mode and each section of the SWC condition was calculated for perceptual attributes that yielded a significant difference between ratings given for stationary and SWC conditions of the VC D and F. The metrics yielding significant correlations were used to investigate participants' association between perceptual attributes and the sound segments in the discussion section. The correlation values between perceptual attributes' ratings and the relevant metric values for stationary and the SWC mode stimuli are given in Tab. 2.

## 4. DISCUSSION

Seven VCs were recorded in a semi-anechoic chamber to use in a listening experiment that utilized a semantic differential technique for determining perceptual attributes that were found relevant in OPQ. Two sets of stimuli, stationary and SWC conditions of seven VCs, were presented to 21 participants in the experiment. Results showed that participants rated stationary and SWC mode stimuli differently in OPQ and in perceptual attributes, indicating that participants focused on different acoustical properties of the sounds in the two conditions.

In the following subsections, we investigate the shift in participants' focus between the two modes by examing the differences between VCs D and F on a percep-

**Table 1**. List of Pearson's correlation and corresponding *p*-values for the selected perceptual attributes and the OPQ ratings.

| Perceptual attribute | Pearson's $r$ (stationary) | Pearson's $r$ (SWC) |
|---|---|---|
| Weak - powerful | $r = 0.53, p < 0.001$ | $r = 0.57, p < 0.001$ |
| Inefficient - efficient | $r = 0.73, p < 0.001$ | $r = 0.67, p < 0.001$ |
| Broken - functional | $r = 0.66, p < 0.001$ | $r = 0.64, p < 0.001$ |
| Annoying - pleasant | $r = 0.53, p < 0.001$ | $r = 0.46, p < 0.001$ |
| Not howling - howling | $r = -0.10, p > 0.05$ | $r = 0.05, p > 0.05$ |

**Table 2**. The ascent, descent, and stationary segments report the correlation values between the perceptual ratings collected in SWC conditions and the metric values calculated over the three sections of the SWC stimuli.

| Perceptual attribute | Section | Metric 1 | Metric 2 |
|---|---|---|---|
| Inefficient - efficient | Stationary mode | T ($95\%, r = -0.77$) | S (mean, $r = -0.90$) |
| | Ascent | N (max, $r = 0.79$) | R (mean, $r = -0.88$) |
| Broken - functional | Stationary mode | T ($95\%, r = -0.77$) | S ($90\%, r = -0.86$) |
| | Ascent | FS (max, $r = -0.79$) | N (max, $r = 0.82$) |
| Annoying - pleasant | Stationary mode | S (mean, $r = -0.99$) | |
| | Ascent | R (max, $r = -0.82$) | |
| | Descent | N (max, $r = 0.88$) | |
| | Stationary segment | S (RMS, $r = -0.89$) | |
| Not howling - howling | Ascent | T (mean, $r = 0.79$) | |
| | Descent | T (RMS, $r = 0.85$) | |

tual attribute basis. For each section, we calculated the psychoacoustic metrics over the time selection of the corresponding section and identified the ones that yielded a significant correlation with at least one of the statistically relevant perceptual attributes. We analyze the data of the metrics as well as the recorded sounds of the VCs D and F to determine the parts of the stimuli that contribute to the participants' judgment on perceptual attribute ratings.

### 4.1 Weak - powerful

None of the psychoacoustic metrics revealed a significant correlation with the weak-powerful attribute, hereafter the power attribute, even though the sound significantly impacted the attribute's ratings. We do not have an explanation for the lack of a relationship between the metrics and the attribute.

However, the literature showed that perceptual attributes of power, powerful-powerless, and loudness, loud-soft, were rated similarly in jury tests [2, 20], and N values correlated significantly with perceived quality and other attributes' ratings, such as annoyance. In the present study, we matched the loudness of the VC recordings for the experiment. If participants related the perceptual attribute of power with N, as they related it to the perceptual attribute of loudness in literature, loudness matching might strip the perception of power from the experimental stimuli. A new experiment should be conducted to assess the impact of N on the weak-powerful word pair by comparing loudness-matched and loudness-unmatched VC sounds to verify or discard our hypothesis on the subject.

### 4.2 Inefficient - efficient

The sound had a significant impact on the inefficient-efficient attribute, hereafter efficiency attribute, in both stationary and SWC condition stimuli. Psychoacoustic metrics T and S correlated significantly with the efficiency attribute's ratings collected in stationary mode. The significant negative correlation between efficiency ratings and T suggests that the perception of a tonal component within the wide-band stationary and noisy VC stimulus was perceived as a cue for inefficiency. The significant correlation between the efficiency attribute and S indicates that participants focused on spectral cues during the stationary sections of the VC sounds. The two metrics correlated significantly with sound quality ratings in the literature that used stationary sections of the VC sounds and were used as the basis of regression models (e.g., [2, 6]).

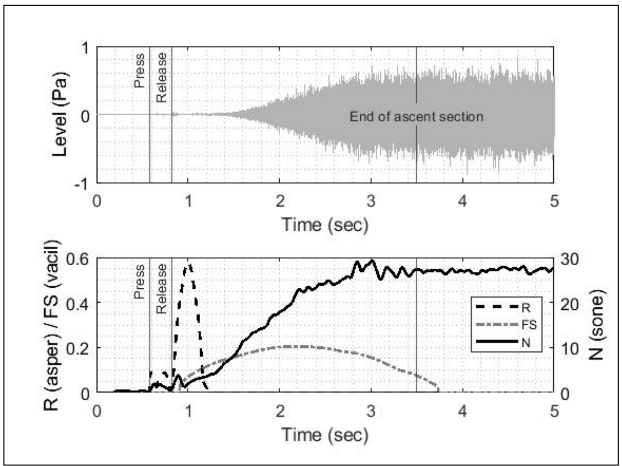


**Figure 3**. The ascent section of the VC D recording is presented in the upper plot with the pressing and releasing moments of the switch denoted. The R, FS, and N values are presented at the bottom.

On the other hand, the efficiency ratings collected in SWC conditions and the psychoacoustic values of maximum N and the average R calculated over the ascent section correlated significantly. The positive correlation with maximum N reveals that matching the N5 of a non-stationary sound, such as SWC stimulus, does not completely eliminate perceptual cues: The maximum loudness values were regarded as indicators of efficiency (see Figures 3 and 4).

R increased briefly after the on/off switches of the VCs were pressed, followed by a second increase coinciding with the switch being released (see Figures 3 and 4). The significant correlation between efficiency and R, together with the time domain observation, suggests that the sound of switching the VCs on/off created an impression on participants, even though the perceived roughness is evoked for a very short period of time. It should be noted that some authors used R values as a parameter in their quality models for VCs in the literature (e.g., [3,5]), while others only used T and S in their models [6].

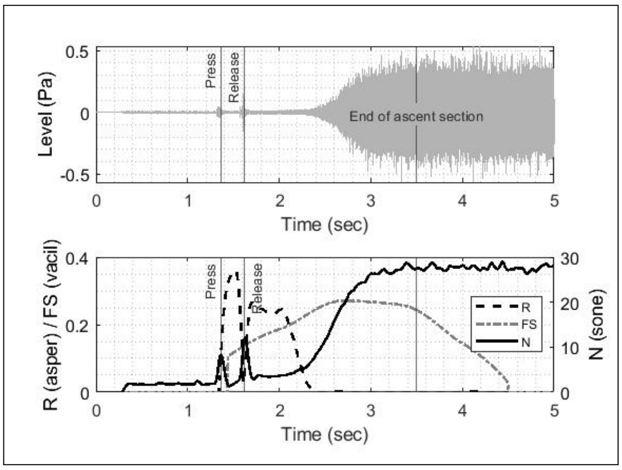


**Figure 4**. The upper plot shows the ascent section of the VC F recording where the pressing and releasing moments of the switch are denoted with vertical straight lines. The bottom plot presents the R, FS, and N values.

### 4.3 Broken - Functional

Similar to efficiency ratings, the sound had a significant effect on the broken-functional ratings, hereafter functionality attribute, in both working modes. The functionality ratings given in stationary mode correlated significantly with T and S. Similar to our observation for the efficiency attribute, the significance of T and S metrics is in line with sound quality literature reporting findings of experiments that used stationary VC sounds as stimuli (e.g., [5]).

For the ascent section, the maximum of FS and N correlated significantly with the functionality ratings. The FS values did not change and were always zero in stationary mode when observed as time series. FS values only increased above zero at ascent and descent sections after the on/off switches were released (see Figures 3 and 4), indicating that the flow rate's increase from zero to maximum generated slow spectro-temporal modulations ($< 20$ Hz) and participants related those with the functionality attribute. The significant correlation between FS and the functionality attribute was shown in the literature for the operation stability of VCs' suction units when tested in different flow rates [18]. The significance of the maximum N suggests that participants relied on the highest loudness reached at the end of the ascent section for the SWC stimuli as they did for the efficiency attribute.

### 4.4 Annoying - pleasant

The annoying-pleasant ratings, hereafter pleasantness attribute, given for stationary mode stimuli negatively correlated with S values, which aligns with the literature [3]. It should be noted that the absolute correlation value between the average S and the pleasantness attribute for the stationary mode condition is the highest among all the metrics correlated for both modes and every section.

The strong relation between the S and the pleasantness was also observed when the SWC condition data were analyzed. The S values calculated for the stationary segment of the SWC mode stimuli, identical to the stationary mode stimuli except for duration, yielded significant correlations with pleasantness ratings given in SWC conditions.

Similar to the efficiency ratings in the ascent section, the R and the pleasantness ratings resulted in a significant negative correlation. The role of maximum R in pleasantness supports our previous observation that participants took the sounds generated by pressing and releasing the switch into account when rating perceptual attributes.

The maximum N correlated significantly with pleasantness when calculated over the descent section. The maximum N in the descent section was reached after the on/off switch was pressed to turn the VCs off. It should be noted that the maximum loudness reached when turning the VCs off is higher than when the VCs were turned on: The sound of turning the VCs off was audible while the VCs were operational and generating loud noise.

### 4.5 Not howling - howling

The perceptual attribute of not howling-howling, hereafter howling attribute, did not result in a significant correlation between ratings given, neither in stationary mode condition nor in SWC conditions when calculated over the stationary segment. The howling attribute yielded a significant positive correlation with T values for ascent and descent sections. The lack of significant correlation between the stationary part of the VC sounds, and the howling attribute makes the howling attribute the only one solely related to the ascent and descent sections of the VC stimuli.

## 5. CONCLUSION

The current study showed that when the vacuum cleaner sounds were presented in single-working-cycle conditions, participants' perceptions of efficiency and functionality were related to the sound they heard when the vacuum cleaners were turned on and immediately afterward. The relationship could be traced with roughness and fluctuation strength values in the time domain for the efficiency and functionality ratings, respectively, for the ascent section. We could conclude that the design of the switches and the shape of the increasing function of the flow rate contribute to the perceptions of efficiency and functionality of the vacuum cleaners, and these attributes' ratings are significantly correlated with overall perceived quality ratings.

The pleasantness attribute was explained by sharpness for both conditions. Additionally, the roughness resulting from the switch sound in the ascent section and the maximum loudness resulting from the switch sound in the descent section correlated highly with the pleasantness ratings, supporting the importance of the sound design of the switches.

The current investigation showed that participants related the tonality of the VCs to the howling attribute during the ascent and the descent sections but not during the stationary parts. When considered together with the non-significant correlation between the howling attribute and overall perceived quality ratings, tonality values resulting from the increase and the decrease of the motor speed and the flow rate of the suction unit do not play a role in overall perceived quality.

The significant correlations between maximum loudness values and different sections of the single-working-cycle condition revealed a limitation that resulted from the difficulty of reducing the impact of loudness on listening experiments when the stimuli are not stationary: The time for the vacuum cleaner to reach its maximum power after being turned on and fully decaying after being turned off is different for each brand used in the study. Hence, loudness matching using an averaging method, such as N5, by applying linear gain was not sufficient to filter out the impact of loudness on perception.

Lastly, the low number of participants and the small differences in overall perceived quality ratings for the vacuum cleaners D and F in two conditions present themselves as the second limitation of the study. Even the t-tests revealed a statistical difference, the effect sizes are rather low. A follow-up experiment could clarify the adequacy of the results.

## 6. REFERENCES


[1] W.-H. Jeon, H. S. Rew, and C.-J. Kim, "Aeroacoustic characteristics and noise reduction of a centrifugal fan for a vacuum cleaner," *KSME International Journal*, vol. 18, pp. 185–192, 2004.

[2] E. Altinsoy, G. Kanca, H. Belek, and A. Senur, "A comparative study on the sound quality of wet-and-dry type vacuum cleaners," in $6^{th}$ *international congress on sound and vibration*, pp. 3079–3086, 1999.

[3] J.-G. Ih, D.-H. Lim, S.-H. Shin, and Y. Park, "Experimental design and assessment of product sound quality: application to a vacuum cleaner," *Noise control engineering journal*, vol. 51, no. 4, pp. 244–252, 2003.

[4] R. Jurc, O. Jiříček, and M. Brothánek, "Methods for the assessment of pleasantness in sound quality," *Noise Control Engineering Journal*, vol. 58, no. 1, pp. 62–66, 2010.

[5] M. E. Altinsoy, "Towards an european sound label for household appliances: Psychoacoustical aspects and challenges," in *Proc. of* $4^{th}$ *International Workshop on Perceptual Quality of Systems. Wien*, pp. 85–90, 2013.

[6] D. Hülsmeier, L. Schell-Majoor, J. Rennies, and S. van de Par, "Perception of sound quality of product sounds a subjective study using a semantic differential," in *Proc. of the International Congress on Noise Control Engineering*, pp. 843–851, 2014.

[7] R. H. Lyon, "Product sound quality-from perception to design," *Sound and vibration*, vol. 37, no. 3, pp. 18–23, 2003.

[8] G. Von Bismarck, "Timbre of steady sounds: A factorial investigation of its verbal attributes," *Acta Acustica united with Acustica*, vol. 30, no. 3, pp. 146–159, 1974.

[9] R. Guski, "Psychological methods for evaluating sound quality and assessing acoustic information," *Acta Acustica united with Acustica*, vol. 83, no. 5, pp. 765–774, 1997.

[10] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and models*, vol. 22. Springer Science & Business Media, 2013.

[11] ISO-532-1, "Acoustics–methods for calculating loudness–part 1: Zwicker method," 2017.

[12] ECMA-418-2, "Psychoacoustic metrics for itt equipment – part 2 (models based on human perception)," 2020.

[13] R. P. Leite, S. Paul, and S. N. Gerges, "A sound quality-based investigation of the hvac system noise of an automobile model," *Applied Acoustics*, vol. 70, no. 4, pp. 636–645, 2009.

[14] J. Y. Jeon, J. You, and H. Y. Chang, "Sound radiation and sound quality characteristics of refrigerator noise in real living environments," *Applied acoustics*, vol. 68, no. 10, pp. 1118–1134, 2007.

[15] J. Kang and M. Zhang, "Semantic differential analysis of the soundscape in urban open public spaces," *Building and environment*, vol. 45, no. 1, pp. 150–157, 2010.

[16] M. Takada, S. Arase, K. Tanaka, and S.-i. Iwamiya, "Economic valuation of the sound quality of noise emitted from vacuum cleaners and hairdryers by conjoint analysis," *Noise Control Engineering Journal*, vol. 57, no. 3, pp. 263–278, 2009.

[17] J. Prezelj and T. Novaković, "Centrifugal fan with inclined blades for vacuum cleaner motor," *Applied acoustics*, vol. 140, pp. 13–23, 2018.

[18] T. Novaković, M. Ogris, and J. Prezelj, "Validating impeller geometry optimization for sound quality based on psychoacoustics metrics," *Applied acoustics*, vol. 157, p. 107013, 2020.

[19] A. Fiebig, "Influence of context effects on sound quality assessments," in *Proc. of EuroNoise*, pp. 2555–2560, 2015.

[20] S. Kuwano, S. Namba, T. Hashimoto, B. Berglund, Z. Da Rui, A. Schick, H. Hoege, and M. Florentine, "Emotional expression of noise: a cross-cultural study," *Journal of sound and vibration*, vol. 151, no. 3, pp. 421–428, 1991.