Baristas' Ear Candy – Audio Cues for Well-Sounding Brews

Sarah Ambros[1], Christoph Reuter[1], Isabella Czedik-Eysenberg, Jian Yang[2], Michael Kuka[3]
[1] *Universität Wien, Institut für Musikwissenschaft, 1090 Wien,*
E-Mail: sarah.ambros@univie.ac.at, christoph.reuter@univie.ac.at, isabella.czedik-eysenberg@univie.ac.at
[2] *Shanghai Conservatory of Music, E-Mail: yangjian@shcmusic.edu.cn*
[3] *Robert Bosch GmbH, E-Mail: michael.kuka@de.bosch.com*

Background

Sounds of products and technical devices are an essential part of their product experience and can regulate the perceived product quality in consumers [1]. To state a few examples: studies show that in refrigerator noises mainly the sounds' loudness and roughness influence the perceived noisiness (e.g. [2]). In vehicle noises lower roughness and sharpness (in combustion engines, e.g. [3]) and higher brightness of the sounds (in electric engines [4]) drive a positive rating. In the case of coffee machines, especially loudness and sharpness of machine sounds are often perceived negatively [e.g. 5].

Aim and research question

The aim of this study is the evaluation of coffee machine sounds on a visual presented rating scale. Results are further compared to ratings of electric engine sounds which were collected in a similar manner [see 4]. The outcome will be picked up in a following experiment, where images of coffee machines are going to be paired with sounds in order to evaluate their audiovisual crossmodal fit.

How do audio features of brewing sounds impact a positive sound experience and the perceived value of coffee machines?

Method and stimuli

Brewing sounds were rated in an online listening task by 73 participants (34♂, 38♀, 1⚥, aged: 19-79, ø 35.4 years) in terms of valence (*Self-Assessment Manikin* [6]) and on a scale from 150 – 1500 € regarding the value/quality of the sounds. These dependent variables valence and value were both implemented graphically.

In the beginning, test subjects adjusted the loudness of their device (based on a -42 LUFS sine wave) and provided information on coffee consumption habits and possibly owned coffee machines. To be able to account for differences in experience, participants provide self-declarations whether they deal with music or audio professionally or not. They further self-estimated a total of 21 personality characteristics on a 5-point Likert scale (see *SEPPO* [7]; successfully applied in [8]). These items were summarized in 3 personality scales regarding extraversion ($\alpha = 0.718$), neuroticism ($\alpha = 0.684$) and psychoticism ($\alpha = 0.219$), as well as one additional dimension for openness to new experiences ($\alpha = 0.651$).

Independent variables of 13 brewing sounds, mainly taken from filter coffee machines from a noise database were set to equalized loudness. They were edited to a length between 16 and 18 seconds and analyzed using Audio Toolbox and MIRtoolbox [9] in MATLAB in terms of 230 audio features. Correlations observed in pilot experiences between sound ratings and spectral fluctuations (spectral flux) as well as brightness (spectral centroid) were taken into account to select the stimuli. Raters' assessments were analyzed in terms of correlations with audio features and group differences.

Results

Assessments of valence of sounds and estimated quality are interrelated ($r = 0.515$; $p < .01$) (Fig. 1).

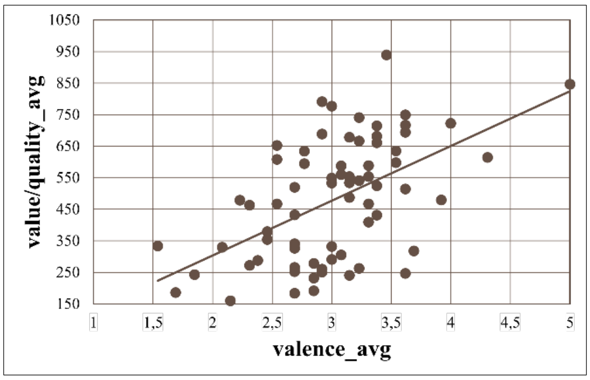

Figure 1: Correlation of valence and value/quality ratings of 13 brewing sounds.

Perceived valence is positively correlated with the strength of the fundamental ($r = 0.555$). Negative associations are linked to greater attack time ($r = -0.615$) and noise fluctuations ($r = -0.596$). Timbral brightness of the stimuli shows a positive correlation with the participants' value/quality ratings ($r = 0.555$), and an even greater correlation with spectral flatness ($r = 0.635$) can be observed. Loudness based on ISO 532-1 (Zwicker) on the other hand reveals negative influences on value/quality ratings ($r = -0.612$) (Fig. 2).

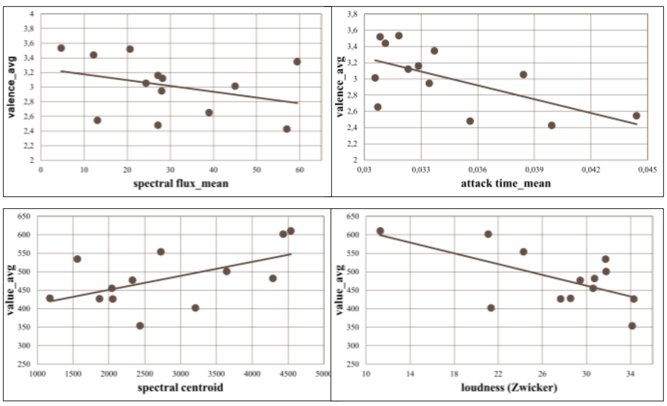

Figure 2: Correlation of valence and value/quality ratings in terms of the sounds' audio features.

It is noticeable that value/quality ratings of tested brewing sounds are generally low and show more statical connections with audio features compared to valence ratings. A possible reason for this observation is that quality rating of sounds and products in particular, might be influenced by a combination of several characteristics and can furthermore be mediated more strongly by demographic data of raters (e.g. status, profession, age).

Differences in terms of personality

After calculation of the participants' expression in regards to the observed personality dimensions, possible group differences in valence and value ratings were examined with t-tests and ANOVAs.

No significant differences can be stated in mean ratings, however when looking at the overall best rated sound (sound20), answers of participants with openness-scores below median value (0) are more widely spread and slightly higher compared to more open listeners (1) ($t_{(73)} = 2.329$; $p < .05$) (Fig. 3). With a view to audio features, this stimulus had the highest level of sounds' brightness as well as the lowest amount of fluctuations.

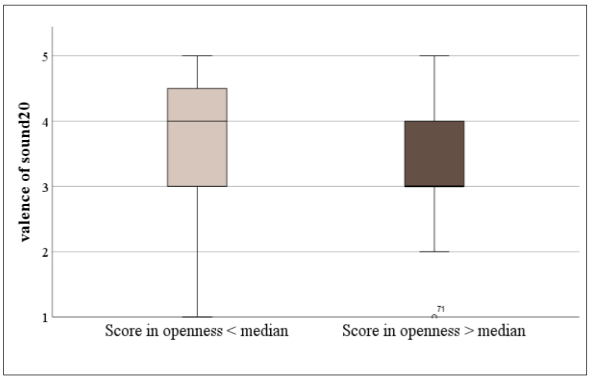

Figure 3: Differences in valence ratings of the overall best rated sound (sound20) regarding participants' scores in openness to new experiences.

Differences in terms of demography and habits

Valence rating of the presented sounds is affected by the gender of participants. Female raters give brewing sounds on average a higher valence rating than male listeners ($t_{(72)} = 2.094$; $p < .05$) (Fig. 4).

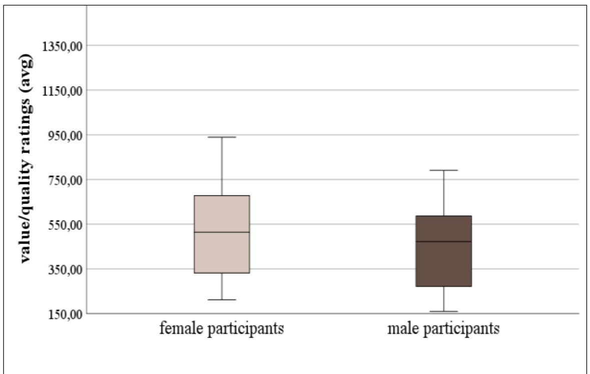

Figure 4: Average value/quality rating for female and male participants.

Further, the individual coffee consumption habits as well as owned coffee machines were used to group the participants into 3 sub-groups: non-coffee-drinkers (n = 13), coffee-drinkers (n = 56) and occasional coffee-drinkers. As only a few participants declared themselves as occasional coffee-drinkers, this group was excluded from group differences analysis.

Overall, there are no significant differences in valence and value/quality ratings with mean valence ratings of 3.14 / 2.99 points and price ratings of 508.79 / 474.32 € in non-coffee consumers vs. coffee-drinkers.

Participants mention overall 4 different coffee machines which are used daily: filter, capsule, automatic, espresso and portafilter coffee machines. Further, a few participants stated owning more than two or three coffee machines. ANOVA show differences in mean value/quality ratings for participants who own one type of coffee machine regarding the owned type of machine ($F_{(5,54)} = 2.881$; $p < .05$). Portafilter coffee machine users (n = 14) provide significantly lower ratings than espresso coffee machine users (n = 8) (Fig. 5).

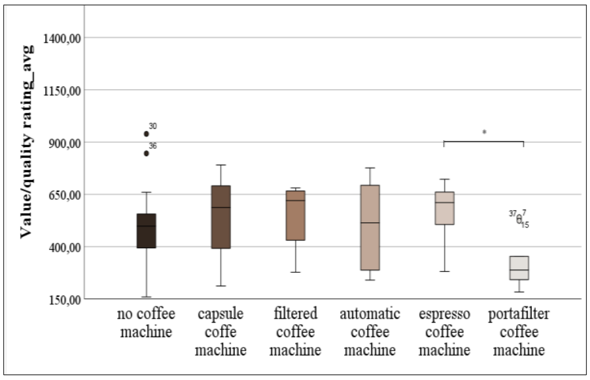

Figure 5: Boxplots of value/quality ratings in terms of different types of coffee machines owned by participants.

Further, the age of participants was tested in terms of correlations with mean valence and value ratings. There is a small effect that with increasing age of raters, the value/quality of sounds decreases ($r = -0.272$; $p < .05$) (Fig. 6).

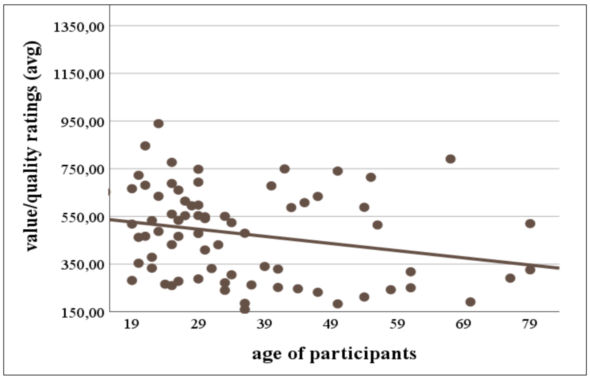

Figure 6: Correlation of value/quality ratings in terms of participants' age.

The experience with music or audio in general (for example due to an occupation or theme-related studies) reveals significant influences on both valence ($t_{(73)} = -2,154$; $p < .05$) and value/quality ratings ($t_{(73)} = -2.006$; $p < .05$) (Fig. 7). Self-declared experienced ears (n = 23) evaluate brewing sounds on average 94 € cheaper than unexperienced listeners (n = 50).

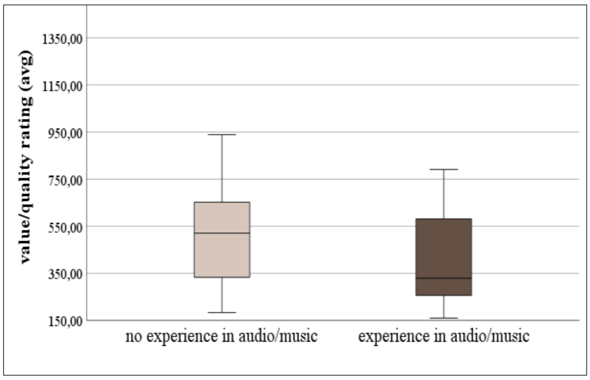

Figure 7: Differences in mean value/quality ratings regarding listeners' self-declared experience in the field of audio/music.

Summary and outlook

The aim of this study was the evaluation of brewing sounds in an online listening task in terms of valence and value/quality of the coffee machines producing the sounds.

Results were analyzed regarding audio features of the presented sounds. This confirmed formerly observed correlations of loudness [2;5] with lower perceived value/quality. On the other hand, increasing brightness in sounds (spectral centroid) was connected to higher value ratings of the sounds, which was also found in noises from other technical sources (e.g. [4]).

With regard to participant-related factors, variance in ratings can be explained mainly regarding the experience of participants with audio or music. Overall, value/quality ratings were low and even lower in experienced ears. In fact, considering most of the presented sounds were taken from machines on the cheaper side of value scale, experienced participants assessments tend to be more in line with reality.

It is planned to take into account the results of this study and former observed connections in similar studies in a follow-up listening experiment on the audiovisual crossmodal fit of products regarding their technical sounds.

References

[1] Hülsmeier, D.; Schell-Majoor, L.; Rennies, J. & van de Par, S. (2014): Perception of sound quality of product sounds. A subjective study using a semantic differential. *Internoise*, Melbourne, 16.-19.11. 2014.

[2] Sato, S.; You, J. & Jeon, J. Y. (2007): Sound quality characteristics of refrigerator noise in real living environments with relation to psychoacoustical and autocorrelation function parameters. *JASA 122/314,* p. 314–325.

[3] Xu, Z.; Xia, X.; Lai, X. & He, Z. (2018): Improvement of interior sound quality for passenger car. *Applied Acoustics 130*, p. 43–51.

[4] Ambros, S.; Kuka, M.; Czedik-Eysenberg, I.; von Georgi, R. & Reuter, C. (2023): Fahrzeuggeräusche in Form und Farbe – Entwicklung eines Crossmodalen Ansatzes zur Erfassung Subjektiver Bewertungen von Elektrifizierten Fahrzeugen. *Fortschritte der Akustik - DAGA 2023, 49. Deutsche Jahrestagung für Akustik*, p. 490–493.

[5] Knöferle, K. (2012): Using Customer Insights to Improve Product Sound Design. *Marketing Review St. Gallen 29/2*, p. 47–53.

[6] Bradley, M.M. & Lang, P.J. (1994): Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *J. of Behavior Therapy & Exp. Psychiatry 25/1*, p. 49–59.

[7] von Georgi, R. & Herr, J.: SEPPO - Short Eysenck Personality Profiler with NEO-PI-R Openness. Unpublished ongoing research project. Berlin Institute of Biomusicology and Empirical Research (BIBER), *Berlin School of Popular Arts*, 2020.

[8] Herr, J., von Georgi, R., & Starcke, K. (2022): Massenmagnet oder Publikumskiller? Eine experimentelle online-Studie zum Einfluss von Hintergrundmusik und Persönlichkeit auf die subjektive Bewertung des präsentierenden Unternehmens eines Messestandes. Zanger, C. (Hrsg.): *Events und Wege aus der Krise.* Wiesbaden: Springer Gabler, p. 139–158.

[9] Lartillot, O., & Toiviainen, P. (2007, September). A Matlab toolbox for musical feature extraction from audio. *International Conference on digital audio effects Vol. 237*, p. 244.