That sounds healthy! Audio and visual frequency differences in brand sound logos modify the perception of food healthfulness

Monin Techawachirakul\textsuperscript{a,b,1}, Abhishek Pathak\textsuperscript{c}, and Gemma Anne Calvert\textsuperscript{a}

\textsuperscript{a} Nanyang Business School, Nanyang Technological University, Singapore

\textsuperscript{b} Department of Psychology, Thammasat University, Thailand

\textsuperscript{c} School of Business, University of Dundee, UK

Abstract

Sonic logos (or brand sound signatures), also termed “sogos”, are a marketing communication tool that brands invest in heavily to signify brand or product benefits to consumers in catchy, non-visual ways. Given the considerable utility of brand sogos, it is surprising that to date there has been scant research into the nature of the specific acoustic features that can be modulated to connote certain traits, including in the current paper, to signify the healthfulness of products within the food category. Our findings as described across three studies revealed that sogos created with higher (vs. lower) frequency were indeed significantly matched with healthy food products (vs. less healthy), while the effect of tempo was neutral. Interestingly, this effect generalizes to high (vs. low) spatial frequency visual stimuli too. The current study contributes to literature on the crossmodal correspondence between acoustic sound clips and expectations of healthfulness. It also advances the theoretical insights into business

\textsuperscript{1} Corresponding author at: Division of Marketing, Nanyang Business School, 48 Nanyang Avenue, 639818, Singapore. E-mail address: Monin001@e.ntu.edu.sg
applications using optimal sogos congruent with visual cues on packaging to connote food healthfulness intuitively to consumers.

**Keywords:** Crossmodal correspondence; Sonic logos; Food healthfulness; Frequency; Tempo; Sensory branding
That sounds healthy! Audio and visual frequency differences in brand sound logos modify the perception of food healthfulness

1. Introduction

Sensory stimuli (e.g., sounds, colours) play a prominent role in branding. A sonic logo (also termed a sogo) is an integral part of branding and is comprised of a short melody created to support the marketing of a specific brand by some semantic association with the unique selling proposition that the brand offers e.g., McDonald’s “I’m loving it” sogo which is crafted in a major key such that it sounds pleasing and playful (Bonde & Hansen, 2013). Compared to other audio assets, sogos tend to be more flexible and adjustable as they can be used across a variety of physical and digital touchpoints to help connect brands to customers at a deep, implicit level in a less cluttered (non-visual) sensory domain (Bonde & Hansen, 2013; Goyal, 2019). These unique properties distinguish sogos from other audible marketing tools, such as background music played in-store or TV advertisements which are not designed necessarily to convey brand identity. Brands thus invest time and financial resources to the creation of these distinctive sound signatures which become a part of their brand identity (e.g., Windows’s start-up chime) and consumers quickly learn the association between these short clips and the brands they represent (Goyal, 2019). Importantly, brand managers and agencies need to ensure that sogos are perceived as congruent with the brand and/or product they represent. Recent research suggests that when music is congruent with the associated food/beverage attributes (e.g., ‘sweet music’ played with the presentation of a sweet meal), participants fixate their eyes on the meal longer, reflecting higher attention, than incongruent combinations (Peng-Li, Byrne, Chan, & Wang, 2020). Congruency between product category and the sound of a brand name can also influence positive
responses towards food products including taste expectations and purchase intention (Fenko, Lotterman, & Galetzka, 2016; Lee & Labroo, 2004; Reber, Schwarz, & Winkielman, 2004). The authors suggest this is because such congruent stimuli enhance the ease of processing, which in turn leads to a more pleasant experience. Thus, achieving this sense of perceptual congruence is important in the context of sogos.

The growth of consumer interest in healthier food choices has urged brands to strive for strategies by which to communicate food healthfulness in competitive markets. While past research on acoustic influences on consumer decision making has mostly focused on how ambient music and background noise influences choice behavior (including the selection of healthy items), similar research on sogos is sparse (e.g., Biswas, Lund, & Szocs, 2019; Mathiesen, Mielby, Byrne, & Wang, 2020). For example, sogos for junk foods will inevitably sound different from sogos for healthy alternatives. Within the food category, creatively crafted sogos have been used to connote food flavours (e.g., vanilla and citrus) (Bronner, Frieler, Bruhn, Hirt, & Piper, 2012), to create brand recognition (Bonde & Hansen, 2013), and to influence willingness-to-buy the associated product (Krishnan, Kellaris, & Aurand, 2012). Given the considerable utility of brand sogos, it is surprising that to date there has been scant research into the nature of the specific acoustic musical features that can be modulated to connote certain traits (e.g., healthfulness). While previous studies have demonstrated that music can influence consumer decision making processes towards healthy foods through underlying mechanisms of morality cues (Huang & Labroo, 2019), music-evoked emotion and arousal (Motoki, Takahashi, Velasco, & Spence, 2022), and visual attention (Peng-Li, Mathiesen, Chan, Byrne, & Wang, 2021), the current research enriches the literature on acoustic influences on consumer perception.
by systematically evaluating the effect of frequency and tempo present in hypothetical brand sogos and the association with food healthfulness.

2. Theoretical background

2.1 Frequency

Frequency is the most critical aspect of sound that makes a commercial sogo unique as it is more discriminable than other musical features (Bonde & Hansen, 2013). In the last decade, the association between the frequency of a sound and the attributes of food products (e.g., taste, crispiness, freshness, harshness) has been investigated extensively (e.g., see Spence et al., 2019 for a review). For example, high (vs. low) frequency sounds have been reliably mapped with sweet (vs. bitter) tastes (e.g., Reinoso-Carvalho, Wang, De Causmaecker, Steenhaut, Van Ee, & Spence, 2016) and crispiness (Zampini & Spence, 2004). Similarly, alcoholic beverages bearing names with high (vs. low) frequency sounds are perceived as harsh (vs. mild) (Pathak, Calvert, & Lim, 2020).

Relevant to the current paper, high (vs. low) frequency sounds present in linguistic stimuli (e.g., brand names) are perceived as healthy (vs. unhealthy) (Motoki, Park, Pathak, & Spence, 2021). Similarly, high (vs. low) frequency sounds have been shown to be linked with small and light objects (Walker, 2016) and healthy (vs. unhealthy) foods are linked with small and light portions (Motoki et al., 2021). Background sounds (e.g., music, ambient sounds) can also influence eating behaviour, such that high (vs. low) frequency sounds often lead consumers to select small (vs. bigger) portion sizes (Lowe & Haws, 2017; Lowe, Ringler, & Haws, 2018). It has also been shown that consumers make healthier (vs. unhealthy) food choices when ambient
THAT SOUNDS HEALTHY (Word count: 5990)

sounds are played in high (vs. low) frequency (Huang & Labroo, 2019). Another study has identified that high (vs. low) frequency is one of the important determinants in creating a healthy (vs. unhealthy) soundtrack (Peng-Li et al., 2021). This appears to be consistent with a recent study revealing that music genre (e.g., classic and jazz) containing high frequency components induces consumers to prefer healthy foods in contrast to rock/metal genres (Motoki et al., 2022). Given this evidence, we predicted that sogos with high (vs. low) frequency sounds would be expected to be more associated with healthier (vs. less healthy) food.

2.2 Tempo

While the existing literature has mostly emphasized the effect of frequency on consumer expectations, this paper also investigates the effect of tempo, which is another important musical component that can also influence food perception (Bronner et al., 2012; Knoferle & Spence, 2012). While there has been no documented report studying solely the effect of tempo on expectations of food healthfulness, recent research has demonstrated that classical and jazz music containing slow tempo affect selections and preferences of healthy foods (Motoki et al., 2022; D. Peng-Li et al., 2021). Peng-Li and colleagues (2021) used a combination of auditory features to compose complex soundtracks including ‘healthy’ (e.g., jazz piece, slow tempo, and high-pitched piano) and ‘unhealthy’ (e.g., fast tempo, low-pitched, and dissonant guitar) soundtracks which are associated with healthy and unhealthy eating respectively. The results showed that the healthy soundtrack influenced participants to select healthy (vs. unhealthy) choices of food items relative to the unhealthy soundtrack. In a similar vein, Motoki and colleagues (2022) had participants listen to music with different music genres and indicated their intention to eat different foods comprised of healthy savoury, healthy sweet, unhealthy savoury, and unhealthy sweet foods. The results showed that classical music which was predominantly
linked with slow tempo pieces led to more preferences for both healthy savoury and healthy sweet foods compared to rock/metal and hip-hop music. Hence, we predicted that sogos with slow (vs. fast) tempo would be expected to be more closely associated with healthier (vs. less healthy) foods.

3. Sogo development

The sogos used in this study were created and manipulated using the Audacity software (https://www.audacityteam.org/). The sogos comprised of novel short melodies without the inclusion of a human voice to rule out the confounding effect of spoken words. The melodies for the sogos were generated using the Random Music Generator online software (https://random-music-generators.herokuapp.com/melody) and comprised of major notes ranging from C4 to B4 (261.63 - 493.88 Hz), moderate tempo (95 beats per minute (bpm)), and a constant rhythm at the eighth note (1/2 beat). The frequency and tempo of each melody was then manipulated. The frequency of melody was shifted by an octave (12 semitones) to create four melodies differing in the frequency range: very low frequency (65.41 - 123.47 Hz), low frequency (130.81 - 146.94 Hz), moderate frequency (261.63 - 493.88 Hz), and high frequency (523.25 – 987.77 Hz). For the tempo manipulation, the tempo of the melody was adjusted to slow tempo (70 bpm), moderate tempo (95 bpm), and fast tempo (145 bpm) (MasterClass, 2020). Once the sogos were created, they were adjusted to be approximately 9 seconds in length with normalized loudness at -20 dB RMS (a sample of sogos is available at https://soundcloud.com/musicclips/10-c4to1b4-95bpm?si=efc3093ee83449399e7cd575625965c).
A pretest was conducted to ensure that the pleasantness of tunes used as sogos\(^2\) in the main studies were controlled. 56 participants (31 males and 25 females) between the ages of 24 to 69 years (\(M = 38.46\) years, \(SD = 11.61\)) were recruited. Using G*Power 3.1.9 (Faul, Erdfelder, Lang, & Buchner, 2007), the sample size had 95% power to detect a medium-sized effect (0.25) in the repeated measure ANOVA. Participants took part in a 4 (frequency: very low, low, moderate, and high) x 3 (tempo: slow, moderate, vs. fast) within-participants pre-test where they rated the pleasantness of sogos on an 11-point Likert scale (from 1 = not pleasing at all to 11 = extremely pleasing). Participants also reported their proficiency in musical training (from 1 = not at all trained to 11 = very well trained), and musical knowledge (from 1 = not at all good to 11 = very good).

Participants rated sogos with very low and moderate frequencies to be significantly different in terms of their pleasantness appeal and thereby these logos were discarded. No difference was found in the pleasantness appeal of sogos created in low and high frequencies and low and fast tempos (frequency: \(F(1,55) = 1.78, p = .19\); tempo: \(F(1,55) = .30, p = .60\)), thereby these logos were retained for the subsequent studies. No effect of the level of musical training nor of the musical knowledge was noted.

4. Method and overview of studies

Participants were presented with hypothetical sogos (study 1 and 2) or visual cues (study 3) followed by images of food products and then asked to select the food that they felt best suited

---

\(^2\) Sogos are referred to tunes (series of musical notes) in this study, the terms ‘tunes’ and ‘sogos’ will be used interchangeably.
for. Each sogo was automatically played one at a time. All studies were programmed on the Qualtrics online survey platform (https://www.qualtrics.com). Participants were native English speakers residing in the USA and recruited from Amazon Mechanical Turk. They were invited to take part in only one of the studies in the present research and were compensated for their time and effort. They were first asked to provide consent and demographical information and then took part in a few practice trials before proceeding to the main task. Using G*Power 3.1.9 (Faul et al., 2007), the sample size (n = 60 for all studies) had 95% power to detect a medium-sized effect (0.5) in the Wilcoxon signed-rank test. The research was approved by the ethics committee of a large northern university in the UK.

5. Study 1

5.1. Participants

60 participants (39 males and 21 females) between the ages of 22 and 63 years (M = 36.67 years, SD = 8.96) took part in the study.

5.2. Stimuli

Sogos selected in the pretest were used as the auditory stimuli. Visual stimuli were selected from the standardized food images of the CROCUFID database (Toet et al., 2019) and F4H Image Collection (Charbonnier, van Meer, van der Laan, Viergever, & Smeets, 2016) to control the colour tones within each food category. The images comprised of three healthy foods (green salad, fruit salad, and tomato and cucumber salad) and three unhealthy foods.
(cheeseburger, French fries, and chocolate cookies) (Bucher, Müller, & Siegrist, 2015; Plasek, Lakner, & Temesi, 2020) (see Appendix 1).

5.3 Design and procedure

A 2 (frequency: high vs. low) x 2 (tempo: slow vs. fast) within-participants design was conducted. Participants were informed that a food company has a range of new products and wanted to select sogos that can convey the healthfulness of its food products. They were instructed to listen to the sogos and then select the food that they felt best matched the sogo. Each sogo was paired with one healthy and unhealthy food image. The pairing of food images was randomized and the position (i.e., left vs. right) was counterbalanced within-participant. The experiment lasted for approximately 8 minutes. To rule out the role of familiarity, participants also rated the familiarity of sogos after the experiment (from 1 = not at all familiar to 11 = very familiar).

5.4. Results and discussion

Data were analysed using SPSS 22.0 for Windows (IBM SPSS Inc., Chicago, IL, USA). As the Shapiro-Wilk Normality test revealed that the null hypotheses of normal population distributions were rejected at $p < .05$, non-parametric statistical methods were used for subsequent data analysis.

For each individual feature of sogos (i.e., frequency and tempo), the Wilcoxon signed-rank test was performed to determine whether participants differ in matching food images with two feature levels. Specifically, the percentage of healthy food selection was compared between low vs. high frequency and slow vs. fast tempo. The results revealed that participants expected
the sogos with high (vs. low) frequency to be more associated with healthy (vs. unhealthy) foods 
\(M_{\text{high}} = 73.89\% , SD = 27.16, M_{\text{low}} = 48.89\%, SD = 32.16; z = 3.64, p < .001, r = .33\), while
they expected sogos with slow (vs. fast) tempo to be more associated with healthy (vs. unhealthy) foods at a marginally significant level 
\(M_{\text{slow}} = 66.11\%, SD = 25.85, M_{\text{fast}} = 56.67\%, SD = 27.82; z = 1.91, p = 0.056, r = .17\) (see Figure 1). No effect of the familiarity of the sogos was found\(^4\). The findings imply that the sogos with high frequency and slow tempo were expected to be more associated with healthy foods (when compared to sogos created in low frequency and fast tempo which were more associated with unhealthy foods).

\[^3\] M is the averaged proportion of trials that participants selected healthy over unhealthy food choices

\[^4\] Though the data was not normally distributed, a repeated measure ANOVA was carried out with familiarity as a covariate. No effect of familiarity was observed (frequency: \(F(1,58) = .98, p = .33\); tempo: \(F(1,58) = .15, p = .70\)
Fig. 1. Percentage of healthy food selection as a function of the frequency (1A, 2A, and 3A) and tempo (1B and 2B). The upper panel represents the responses towards sogos across food categories (Study 1). The middle panel represents the comparison of responses towards sogos within each food category (Study 2). The lower panel illustrates the responses towards grating images across food categories (Study 3). Error bars represent the standard error of the means.

6. Study 2

While Study 1 provided evidence for an association between sogos with healthy vs. unhealthy categories of foods, it is plausible that the category of food (e.g., salad vs. burger) possibly influenced the results. We, therefore, wanted to extend these findings to a similar food category by varying the healthfulness appeal (e.g., by ingredients), such as soy burgers which are perceived as healthier than beef burgers (Motoki et al., 2021). Thus, Study 2 was conducted to test whether the association between sogos and healthfulness extends to foods within the same food category (e.g., healthy burger vs. unhealthy burger or healthy vs. unhealthy salad). If this is the case, the significant association between high (vs. low) frequency and slow (vs. fast) tempo with healthy (vs. less healthy) foods should be observed. As the different colour tones between food categories in Study 1 might have affected the results, in Study 2, colour tones between food choices were matched.

6.1. Participants

59 participants (28 males and 31 females) between the ages of 20 to 68 years ($M = 41.17$ years, $SD = 11.37$) took part in the study.

6.2. Stimuli
Sogos selected in the pretest were used again here as the auditory stimuli. Visual stimuli were selected from freely available resources online. The images comprised of two variations of burger (soy burger vs. bacon burger) and two variations of salad bowl (spring salad vs. spring salad with fries). For all images, special care was taken to ascertain similar backgrounds, viewpoints, and colour tones to minimize any confounds between the images. To create the unhealthy version of a salad, images of fries were superimposed on the salad image to create an unhealthy version (since fries are perceived as unhealthier; Bucher et al., 2015). All images were presented at a resolution of 300 x 300 px (see Appendix 2).

A pretest was conducted to verify the perceived healthfulness of the selected images. 58 participants (23 males, 33 females, and 2 with unspecified gender) between the ages of 26 to 70 years ($M = 44.33$ years, $SD = 12.15$) rated the images on perceived healthfulness on a visual analog scale (VAS) from 0 (not at all healthy) to 100 (very healthy). As expected, an independent samples $t$-test, using SPSS 22.0 for Windows (IBM SPSS Inc., Chicago, IL, USA) revealed that soy burger and spring salad were perceived as healthier than spring salad with fries and bacon burger (soy burger: $M = 62.17$, $SD = 22.80$; bacon burger: $M = 14.36$, $SD = 15.02$; $t(57) = 13.03$, $p < .001$, $d = 1.71$; spring salad: $M = 89.90$, $SD = 11.49$; spring salad with fries: $M = 41.64$, $SD = 25.90$; $t(57) = 14.04$, $p < .001$, $d = 1.84$).

6.3. Design and procedure

A 2 (frequency: high vs. low) x 2 (tempo: slow vs. fast) within-participants design similar to Study 1 was conducted. Participants were presented with a pair of healthier and less healthy foods (e.g., soy burger vs. bacon burger, spring salad vs. spring salad with fries) and were asked
to select the food that they felt best matched the sogo. Subsequently, the data were analysed as described above in Study 1.

6.4. Results and discussion

The Wilcoxon signed-rank test was performed to determine whether participants differ in matching food images with two feature levels. Specifically, the percentage of healthier food selections was compared between sogos with low vs. high frequency and slow vs. fast tempo. The results revealed that participants expected the sogos with high (vs. low) frequency to be more associated with the healthy (vs. less healthy) foods ($M_{\text{high}} = 69.92\%, SD = 28.91$, $M_{\text{low}} = 47.03\%, SD = 31.17$; $z = 3.73, p < .001, r = .34$); no effect of the tempo or type of the food category was observed (tempo: $M_{\text{slow}} = 57.63\%, SD = 32.58$, $M_{\text{fast}} = 57.20\%, SD = 25.87$; $z = .03, p = .97$; type of food category (i.e. burger vs. salad): $F(1, 116) = .34, p = .56$; burger: $M = 57.20\%, SD = 24.59$; salad: $M = 59.75\%, SD = 22.75$).

Data was also analyzed separately for both the categories of food (i.e., burger and salad), which revealed similar results. Participants expected sogos with high (vs. low) frequency to be more associated with healthier (vs. less healthy) food (burger: frequency: $M_{\text{high}} = 68.64\%, SD = 34.61$, $M_{\text{low}} = 45.76\%, SD = 35.10$; $z = 3.17, p = .002, r = .29$; salad: frequency: $M_{\text{high}} = 71.19\%, SD = 33.74$, $M_{\text{low}} = 48.31\%, SD = 37.10$; $z = 2.93, p = .003, r = .27$); no effect of tempo was observed (burger: $M_{\text{slow}} = 59.32\%, SD = 35.32$, $M_{\text{fast}} = 55.08\%, SD = 37.94$; $z = .58, p = .57$; salad: $M_{\text{slow}} = 55.93\%, SD = 36.06$, $M_{\text{fast}} = 59.32\%, SD = 25.38$; $z = .85, p = .40$).

---

5 Though the data was not normally distributed, a mixed repeated measure ANOVA was used to test the effect of the food category (burger vs. salad).
7. Study 3

In addition to acoustic frequency, in a follow-up study, we investigated whether the frequency effects observed in the acoustic domain would generalize to the visual modality (specifically, high versus low frequency stimuli as widely used in the psychophysics literature; Landy, 2006). These might provide some insight as to the characteristics of striped pack designs on healthy foods, e.g., higher frequency gratings may connote healthfulness to a greater extent than low frequency gratings. By analogy, the graphical representation of sound waves are simple sinusoidal (sine) waveforms. The sine wave represents the area of compression and rarefaction of medium molecules (e.g., air) occurring while the sounds are transmitted. This resembles a visual image known as a sine wave grating (stripes of black and white) where the bands represent the low and high intensity of light and its frequency is also measured in terms of the number of cycles per visual degree (higher cycles indicate higher spatial frequency: see Figure 2). In line with the crossmodal effect, participants automatically map high (vs. low) frequency musical tones with high (low) spatial frequency images (Evans & Triesman, 2010). Considering our findings on the correspondence between high frequency sogos and food healthfulness, we hypothesized that if the crossmodal association between high-frequency sounds and healthfulness holds true, this association should also be observed in the presence of high frequency images. We, therefore, used sine-wave grating images (visual) as proxies for the acoustic stimuli in studies 1 and 2 and hypothesized that high (vs. low) spatial frequency grating images would be associated with healthy (vs. unhealthy) foods.
Fig. 2. Sine-wave gratings and measurement of frequency. The frequency of sine-wave gratings is measured in terms of the number of cycles per visual degree. (a) Low spatial frequency (one cycle per degree) (b) High spatial frequency (two cycles per degree) (Kalloniatis & Luu, 2007). Source: http://webvision.med.utah.edu/

7.1. Participants

58 participants (32 males and 26 females) between the ages of 23 to 77 years ($M = 44.10$ years, $SD = 13.76$) took part in the study. Data of one participant who spent over 10 minutes to complete the task was excluded.

7.2. Stimuli

The visual stimuli were black and white sine-wave gratings created using PsychoPy 2 (Peirce et al., 2019). The grating had bars oriented at $0^\circ$, $45^\circ$, $90^\circ$, and $135^\circ$, moderate contrast of 1, and resolution of 280 x 280 px. The spatial frequencies were manipulated at high (12 cycles/degree) and low (4 cycles/degree) frequency (see Appendix 3). Food images comprised of healthy foods (salad, fruit salad, salad with shrimps, and tuna salad) and unhealthy foods (French fries, cheeseburger, brownies and donuts) and were selected from the F4H Image Collection (Charbonnier et al., 2016). Food images were presented at a resolution of 350 x 240 px and were randomly paired and counterbalanced (i.e., left vs. right position) for each participant.
A pre-test was conducted to verify the association of high (vs. low) frequency sogos and high (vs. low) frequency grating images. 60 participants (33 males, 26 females, and 1 with unspecified gender) between 23 – 75 years old (M = 41.53 years, SD = 12.95) were instructed to select a grating image that they felt best matched the sogo. Data of two participants who spent over 10 minutes to complete the task were excluded. As predicted, an independent samples t-test, using SPSS 22.0 for Windows (IBM SPSS Inc., Chicago, IL, USA), revealed that high (vs. low) frequency sogos were associated with high (vs. low) frequency grating images (M_{high-freq} = 75.86\%, SD = 29.97, M_{low-freq} = 50.86\%, SD = 33.10; z = 3.09, p = .002, r = .29).

7.3. Design and procedure

The study utilized a one-way, two-level (high vs. low frequency gratings) within-participant design. Similar to study 1, participants were presented a grating image and a pair of food images and were instructed to select the food that they felt best matched the grating image (a total of eight trials).

7.4. Results and discussion

Data were analysed using SPSS 22.0 for Windows (IBM SPSS Inc., Chicago, IL, USA). As the Shapiro-Wilk Normality test revealed that the null hypotheses of normal population distributions were rejected at p < .05, non-parametric statistical methods were used for subsequent data analysis.

The Wilcoxon signed-rank test was performed to determine whether participants differ in matching food images with low and high frequency. Specifically, the percentage of healthy food selection was compared between images with low vs. high spatial frequency. As predicted, the

---

6 M is the averaged proportion of trials that participants selected high over low frequency grating images
results revealed that participants expected high (vs. low) frequency grating images to be associated with healthy (vs. unhealthy) foods ($M_{high-freq} = 35.09\%, SD = 32.68$, $M_{low-freq} = 24.12\%, SD = 30.24$; $z = 2.13$, $p = 0.03$; $r = 0.20$). The results confirm the generalization of the acoustic frequency effect to visual cues and association with food healthfulness. The implication of this novel finding will be discussed in the next section.

8. General discussion

The primary purpose of the study was to examine whether sogos composed with high (vs. low) frequency and with slow (vs. fast) tempo would be expected to be associated with healthy (vs. unhealthy) foods by systematically manipulating the sound frequency and tempo. The findings partially support the hypotheses. That is, the robust effect of the high (vs. low) frequency sogos on the perception of food healthfulness was observed (S1 and 2), while the effect of slow (vs. fast) tempo was at a marginal level (S1). The current study thus provides initial evidence of the generalization of the frequency effect in the visual domain whereby high frequency grating images are also more associated with healthy foods (S3).

8.1. Sound frequency-healthfulness association

Our findings are in line with previous research suggesting a linkage between high frequency sounds and healthy food products (Motoki et al., 2021; Motoki et al., 2022; Peng-Li et al., 2021). For example, Peng-Li et al., 2021 demonstrated that listening to a ‘healthy’ (vs. ‘unhealthy’) soundtrack comprised of high (low) frequency, slow (fast) tempo, and other features guided participants’ attention towards healthier foods, and lead to healthier food choice.
selections. Our findings further support the observation that frequency itself can potentially connote the attribute of healthfulness. Specifically, Study 1 demonstrated this phenomenon using two broad types of food categories (healthy vs. unhealthy foods) and the results revealed that sogos with high (vs. low) frequency and slow (vs. fast) tempo were expected to be more closely associated with healthy (vs. unhealthy) food products. In Study 2, we extended these findings using foods within the same category (healthy vs. less healthy burgers, healthy vs. less healthy salads) and only the effect of frequency was detected. Given this, the association between frequency and healthy foods can generalize from healthy vs. unhealthy food categories to varying healthfulness appeal within a food category. In other words, sogos with high (vs. low) frequency are associated with healthy (vs. less healthier) foods, irrespective of the food category.

However, recent studies have revealed a boundary condition between tastes/flavors in the frequency effect and showed that listening to music genres comprised of higher frequencies (e.g., classic and jazz) evoked positive valence which in turn appears to increase preferences for sweet and healthier savoury foods compared to rock/metal music (Motoki et al., 2022). Furthermore, Motoki and colleagues (2021) demonstrated that hypothetical brand names containing higher (vs. lower) frequency sounds were rated to be more appropriate for healthy savoury foods while there was no significant difference in the sweet food category. It should be noted that food stimuli used in this study contained savoury and sweet foods in both healthy and unhealthy food categories which is not the case in our current study. This issue will be discussed in a later section.

In this study, the automatic association between high frequency stimuli and food healthfulness can possibly be explained by the shared semantic association between the features of frequency and food stimuli. We assume that high frequency sogos and healthfulness are
preferably matched because both stimuli signify smallness and lightness, while low frequency
and less healthful indicators are preferably associated because both features connote largeness
and heaviness (e.g., Lowe & Haws, 2017; Motoki et al., 2021). Additionally, we also provided
evidence that sogos used here are matched in terms of tune pleasantness (in the process of
fictitious sogo development), and the covariate effect of tune familiarity was not observed (Study
1); hence, it is less likely that our findings will be explained by tune pleasantness and familiarity.

8.2. Sound tempo-healthfulness associations

Although there has been evidence that music consisting of slow tempo increased choice
selection and preferences towards healthy food items (Motoki et al., 2022; Peng-Li et al., 2021),
no significant difference between the matching of healthy foods with slow versus fast tempo
sogos was observed here. It is possible that our sogo stimuli were less complex than music
stimuli used in the prior studies. Moreover, when the tempo is manipulated solely, its effect on
consumer responses might differ from when it is mixed with other musical attributes. Perhaps the
other auditory parameters (e.g., high/low frequency) might override the tempo effect on healthy
food preferences in the previous studies using music with greater compositional complexity
(Motoki et al., 2022; Peng-Li et al., 2021). Furthermore, considering the tempo itself, previous
research has demonstrated that tones with increasing tempo led to increasing movement speed
(Küssner et al., 2014). Much in line with research on food-related behaviours, participants
chewed more rapidly when high (vs. low/no) tempo music was played (Roballey et al., 1985).
Participants exhibited longer (vs shorter) eating times when music was played at a slower (vs
faster) tempo (Mathiesen et al., 2020). Given this, it is plausible to suggest that tempo is more
related to consumption rate (see Spence et al., 2019 for a review) rather than food perception and
this might explain why solely varying tempo (while other parameters remain constant) does not differently influence association with healthy vs unhealthy foods.

### 8.3. Visual frequency-healthfulness associations

While the previous literature has provided evidence of the crossmodal correspondence between high (vs. low) frequency musical tones with high (low) spatial frequency images (Evans & Triesman, 2010), our findings extend this correspondence effect to high frequency-healthfulness association in both auditory and visual stimuli. Specifically, in Study 3, we extended the frequency-healthy association to visual stimuli and the results confirmed that similar to high frequency sounds, high (vs. low) frequency images are similarly matched with healthy (vs. unhealthy) foods. Given that grating images with high spatial frequency correspond to smaller stripes and finer detail, whereas low spatial frequencies represent larger stripes and encode coarseness (Bar, 2004), it is possible that the high (vs. low) frequency gratings connote the trait of healthfulness through the shared semantic meaning in a similar way to auditory stimuli. Collectively, high (vs. low) frequency stimuli, both auditory and visual are crossmodally correspondent with healthy (vs. less healthy) attributes in food products.

### 9. Theoretical and managerial contributions

By linking frequency to healthfulness, this paper offers two specific contributions. First, the current study adds evidence to the extant literature on sound-healthfulness associations by providing evidence that sogos can shape the perception of healthy food amongst consumers by manipulation of stimulus frequency. Moreover, these findings illustrate that the sound frequency
effect generalizes to visual stimuli by showing that high (vs low) frequency images are also associated with healthfulness.

Second, the findings provide a managerial contribution to brand managers and their agencies who are seeking a novel sensory branding strategy for healthy food products. Specifically, we have advanced the application to the marketing of cross-modal correspondences such that high frequency sogos (and potentially slow tempo) can enhance the healthy appeal of foods. Moreover, the effectiveness of sogos can be enhanced by congruent packaging or visual logos using high frequency images (e.g., differing design of stripes or bars) to lead consumers into making healthier choices without having to evaluate complex information often contained in the nutritional labels. Thus, incorporating the healthy sound logos and high frequency visual cues in food brands could boost healthy food choices at an implicit level.

10. Limitations and future research directions

Firstly, although the findings imply how the inclusion of high frequency components in sogos could signify the healthfulness of food products, future research could further investigate whether and to what extent the effect generalizes to actual sogos which involves more compositional complexity (e.g., instrumental timbre, intensity fade up/down, pitch ascending/descending) (e.g., Mas et al., 2020). The fictitious sogos used in the current study (which comprise sound clips modified in terms of their frequency and tempo) were created with less complexity in order to control for potential confounds associated with naturalistic stimuli in the first instance. Future research could thus investigate the effect of sogos crafted with more complex parameters to gain further insights as to how combinations of music parameters can
convey desired product attributes (including beyond the trait of healthfulness). Moreover, as the sogos used in this paper were manipulated in an experimental setting, adopting actual sogos would perhaps increase the ecological validity of the findings.

Secondly, although high frequency is found to be associated with healthfulness, it is also linked with sweetness (e.g., Knoferle & Spence, 2012). Given that sounds differently influence the perception of sweet and savory foods (Motoki et al., 2021, 2022), it is interesting that somewhat differential results were obtained here. In study 1 and 3 we included both savoury and sweet foods in both the healthy and unhealthy categories. Thus, it should be noted that our finding of high frequency-healthfulness association might be partially influenced by the inclusion of sweet foods in the healthy category, and the low frequency-unhealthfulness association might be also affected by the inclusion of savoury foods in the unhealthy category. In addition, in Study 2 all food stimuli were savoury; hence, the frequency effect found here might be partly affected by the saltiness of the foods. However, it is important to note that since the sounds used in previous research were composed with more complex features than the current study, the music genre (with compositional complexity), instead of frequency solely, may serve as a determinant for taste and healthfulness evaluations (Motoki et al., 2022). Nonetheless, before firm conclusions can be drawn, further investigations are required. For example, subsequent research should consider examining whether the frequency solely affects evaluations of food healthfulness combined with different tastes.

Thirdly, an intriguing clue to the biological basis of this effect comes from research showing how noises alter the taste perception of food being consumed (Yan & Dando, 2015). The authors explain that while the chorda tympani nerve carries taste information from the anterior portion of the tongue to the brain, it crosses the tympanic membrane of the middle ear.
Loud noises temporarily disturb the transition of taste signals while they are crossing through the middle ear such that sweet taste signals become weakened and umami tastes become strengthened, leading to the altered perception of food tastes (Yan & Dando, 2015). Based on this observation, an interesting topic for future research would be to investigate the precise neural mechanisms underlying the association between frequency and perceptions of healthfulness. Given that they are possibly linked by some shared semantic association, the study of the neural correlates of semantic associations and conceptual similarity, in regions such as the anterior temporal lobe (Jackson, Hoffman, Pobric, & Lambon Ralph, 2015) would possibly help us gain more insight into how acoustic stimuli are crossmodally associated with such abstract concept as healthfulness in the brain.

Fourthly, although at this point, we have introduced the new finding of visual frequency-healthfulness association, further research is needed to investigate whether adding a design of a high frequency grating image would enhance (i.e., the additive effect) the appeal of the healthfulness of a product. Further, our preliminary findings have yet to illuminate the boundary conditions under which these audio-visual frequency-healthfulness effects occur. Thus, additional research is needed to investigate the moderating role of health concern/motivation on the implicit association between sogos and perceived healthfulness. Lastly, the current paper used only the piano to create sogos, future research could explore the role of timbre or other musical instruments on the perception of food healthfulness.
References


Effects of ambient music and background noise on food sales. Journal of the Academy of

orientation: Rudiments of quantitative research integrating branding theory, social
semiotics and music psychology. SoundEffects-An Interdisciplinary Journal of Sound and
Sound Experience, 3(1-2), 112-135.

Bronner, K., Frieler, K., Bruhn, H., Hirt, R., & Piper, D. (2012). What is the sound of citrus?
Research on the correspondences between the perception of sound and flavour.

scores and subjective lay evaluations in comparison. Appetite, 95, 408-414.
https://doi.org/10.1016/j.appet.2015.08.005

Appetite, 96, 166-173. https://doi.org/https://doi.org/10.1016/j.appet.2015.08.041

Evans, K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory
features. Journal of vision, 10. https://doi.org/10.1167/10.1.6


THAT SOUNDS HEALTHY (Word count: 5990)

https://doi.org/10.1509/jmr.14.0300


MasterClass. (2020, November 8). *Music 101: What is tempo? How is tempo used in music?*
https://www.masterclass.com/articles/music-101-what-is-tempo-how-is-tempo-used-in-music#what-are-the-basic-tempo-markings


https://doi.org/10.1016/j.foodqual.2020.104157

doi:https://doi.org/10.1016/j.foodqual.2021.104380


Appendix 1

Food images used in Study 1

Healthy foods:

- Green salad
- Fruit salad
- Cucumber and tomato salad

Unhealthy foods:

- Cheeseburger
- French fries
- Chocolate cookies
Appendix 2

Food images used in Study 2

**Burger:**

Healthy burger (soy burger)       Less healthy burger (bacon burger)

![Burger Images](image1.png)

**Salad bowl:**

Healthy salad (spring salad)       Less healthy salad (salad with fries)

![Salad Images](image2.png)
Appendix 3

Visual stimuli used in Study 3

High-frequency sine-wave gratings (spatial frequency = 12)

Low-frequency sine-wave gratings (spatial frequency = 4)
Healthy foods:
- Green salad
- Fruit salad
- Salad with shrimps
- Tuna salad

Unhealthy foods:
- French fries
- Brownies
- Cheeseburgers
- Donuts