Sound quality assessment of sewing machines

James J. Chatterley[a)], Jonathan D. Blotter[b)], Scott D. Sommerfeldt[c)] and Timothy W. Leishman[d)]

(Received 2005 May 12; revised 2006 March 23; accepted 2006 March 24)

This paper presents the sound quality analysis of six sewing machines. The machines range in consumer market segments from entry-level, thru mid-level machines, to high-end computer controlled units. Two brands at the three levels are evaluated and compared. The two methods used to determine the sound quality of these machines are jury based listening tests and quantitative sound quality metrics. The details of these methods are presented and discussed. Sound quality results are presented and suggest that metrics can successfully give an indication of customer preference. The results also reaffirm that sound quality analysis can be useful in product design.

Primary subject classification: 79.2; Secondary subject classification: 12.4

1 INTRODUCTION

Historically, good sound quality was often considered to be synonymous with quiet sounds[1]. Although this classical approach of reducing the overall noise level has improved many products and industrial processes, some products, even though they are relatively quiet, are often unappealing or even annoying[2]. Therefore, as the area of sound quality has developed over the last two decades, the concept and definition of sound quality have also changed. Blauert and Jekosch[3] define sound quality as "the adequacy of a sound in the context of a specific technical goal and/or task." The term "compatibility" has also been used in this context, especially with regard to sounds accompanying actions of product users[4]. With this definition in mind, sound quality now represents the "sensory pleasantness" or appropriateness of the sound, which is often expressed as a combination of various metrics, such as the perceived loudness, roughness, sharpness, pitch and others[5]. The 2003 July-August issue of the Noise Control Engineering Journal is completely dedicated to this topic of sound quality[6]. The difference in definition contrasts the one-dimensional approach initially used to the current multi-faceted technique that includes aspects of psychology and anatomy, as well as the physical parameters engineers and scientists are accustomed to using. Although consumers may not make decisions based exclusively on sound, they often make subconscious decisions as to the quality of a product based on the perceived quality of the sound it produces. The important point is that consumers are not necessarily focused on the product sound, but that the product sound is a carrier of information for them. They certainly would prefer a pleasant sound to an unpleasant one, but even more so, they want the "sound of quality"[7]

This paper presents an analysis of six consumer sewing machines. The objective of the research was to determine consumer preference for the sound emitted from each machine, and then to investigate sound quality metrics to determine if an objective metric could reproduce consumer preferences. The hypothesis of this research was that the perceived machine sound quality follows the expected tier pattern where the high quality sounds are associated with the high-end machines, with sound quality stepping down with each corresponding level of machine tier. The sound quality analysis approach presented in this paper is similar to the approach outlined by Bowen and Lyon[8]. The process of determining the relative sound quality of these machines includes the following:

1. Development of the sewing sounds data base including such sounds as:
   a. Variation in speed and stitch pattern.
   b. Variation in machine isolation from the sewing table, operating at medium speed.
   c. Digital modification of the frequency spectrum.
2. Individual machine sound quality assessment from both juror listening tests and mathematical metrics.
3. Determination of most acoustically pleasant sewing machine.
4. Ranking of machines in order of perceived sound quality.

This paper presents the research conducted, beginning with an overview of the development of the sewing sound database. The sound quality assessment, including both the jury listening tests and the mathematical sound quality metrics is discussed. Finally, the sound quality results are presented, along with subsequent design implications.

2 SEWING SOUND DATA BASE DEVELOPMENT

Sounds from each of the six sewing machines, at three operating speeds and four stitches, as well as start up sounds were collected. The reason for investigating multiple stitch conditions is that different stepper motors are in operation for various stitches. Variations in speed allowed the investigation

---

[a)] *Brigham Young University, Dept. of Mechanical Engineering, Provo, UT 84602, E-mail: james.j.chatterley@lmco.com*

[b)] *Brigham Young University, Dept. of Mechanical Engineering, Provo, UT 84602, E-mail: jblotter@byu.edu*

[c)] *Brigham Young University, Dept. of Physics and Astronomy, Provo, UT 84602, E-mail: scott_sommerfeldt@byu.edu*

[d)] *Brigham Young University, Dept. of Physics and Astronomy, Provo, UT 84602, E-mail: tim_leishman@byu.edu*

of sound as a function of speed. The sound data were collected in stereo using a Sony digital audio tape (DAT) recorder with two integrated-circuit piezoelectric (ICP) microphones and accompanying pre-amplifiers. One-minute sound clips at each of the seventy-two operating conditions were collected. The sound files were then digitally transferred from the tapes to the computer for processing, where shorter sound bytes were cut from the one-minute clips for juror listening tests. The test set up is illustrated in the schematic and photo in Figs. 1 (a&b).

The acoustic editing software program used was Sound Forge 6.0, a commercial software package by Sonic Foundry. Digital modifications to the raw sounds were created, focusing on different attributes of the sounds that could be physically realized. These digital modifications include reduction in the magnitude of sections of the frequency content, the addition of fixed frequency sine waves and the digital removal of clicks and taps. The reductions in the magnitude of sections of the frequency content targeted either high frequencies, above 1 kHz, or low frequencies, below 500 Hz. The high frequency reduction represents an improvement in the passive noise control and the reduction in low frequency content represents a change in the internal components of the machine. Addition of fixed frequency sine waves represents the addition of masking tones. The removal of clicks and taps also represents a change in internal components. Sounds were also recorded which include the addition of foam blocks between the table and the machine to semi isolate the machine from the table. This was done to approximate new foot assemblies.

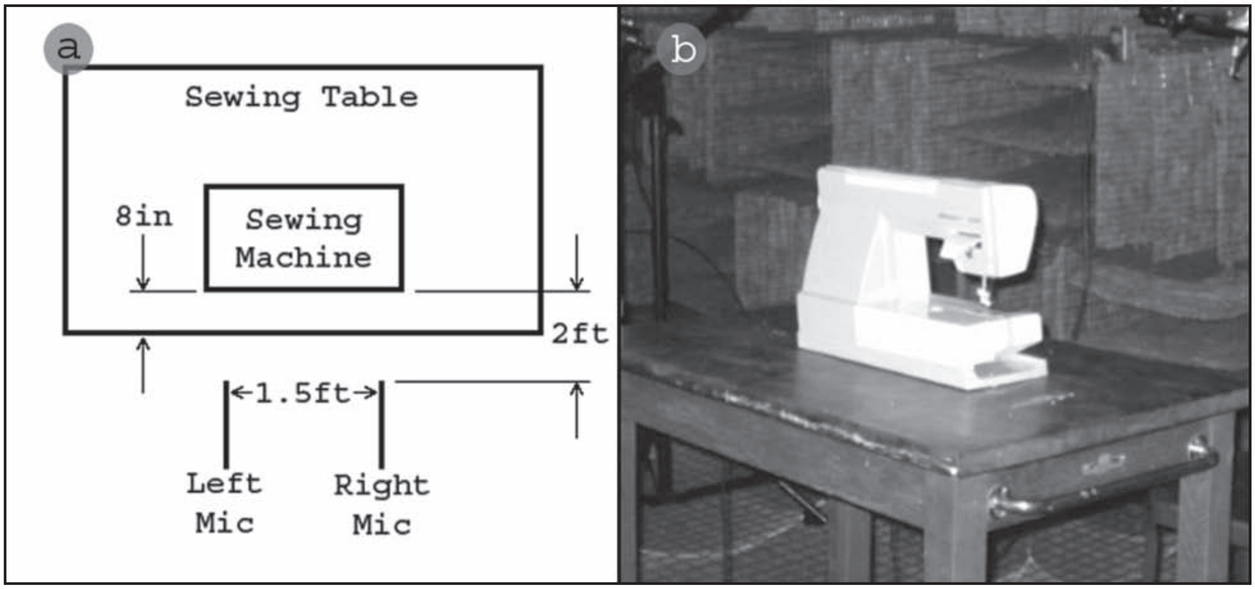
3 SOUND QUALITY ASSESSMENT

To assure that our results were as accurate as possible, both the jury listening test method, as well as mathematical metrics, were used to determine the sound quality of the six sewing machines evaluated. This section reviews both methods briefly to help familiarize the reader with the techniques used in sound quality assessments.

3.1 Jury Listening Tests

The design of the jury listening questionnaire is vital to obtaining meaningful jury results. To aid in the development of the questionnaire, initial jury tests were conducted with a relatively small group of jurors. The results of this initial test were used to fine tune the main jury tests. For example, it was discovered from the preliminary jury testing that some of the physical and digital acoustic modifications produced sounds so close to the original that it was impossible for the jurors to distinguish between the two sounds. Therefore, several sounds were discarded from the study, including those where the bobbin assembly had been removed, the addition of masking tones, and sounds where a click or tap had been digitally removed.

Four separate compact discs were created with each one investigating different sounds. The stitches and speeds compared on each compact disc are listed in Table 1. The jurors listened to multiple stitches, as each stitch requires different stepper motors to operate. Startup sounds were also evaluated to determine if the ramp up and registration of the stepper motors on the computer-controlled machines were acoustically pleasant. Each compact disc contained 42 questions, in which



a) Schematic of test setup b) Photograph of test setup

Fig. 1. Data acquisition setup.

Table 1 - Jury survey CD description

	CD 1	CD 2	CD 3	CD 4
Stitch used in comparison	Reinforced straight stitch (two stepper motors running)	Zigzag stitch (three stepper motors running)	Straight stitch (one stepper motor running)	Machine startup
Additional comparisons	Straight stitch at high speed with modified spectral content	Straight stitch with machine isolated and not isolated from table	Straight stitch at medium speed with modified spectral content	Straight stitch at low speed with modified spectral content

the juror was asked to select one of two sounds and then asked to comment as to why they found it to be more pleasant. The sound clips used for the juror test were six seconds in length, with an approximately 1 second time interval between the two sounds. This afforded enough time for the juror to determine the pleasantness of the sound, yet not so long as to become distracting. Each combination of the machines was presented exactly the same number of times to the jurors, in order to provide consistent data.

Each juror listened to a single compact disc over Sony MDR 7505 studio headphones which provided a clear, high quality, flat reproduction of the sound, as it is imperative that the sounds the jurors listen to are as close to the original sound as possible. For additional discussion on the benefits of headphones in sound quality assessments, the reader is directed to Ref. 4. Figure 2 gives an example of the jurors in the listening environment, where they had minimal distractions, which allowed them to focus on the selected sounds.

Fig. 2. Jurors participating in listening survey.

As the consumer base of high-end sewing machines is comprised of middle-aged women, with moderate to advanced sewing skills, these were the desired demographics of our jury. The demographics of the jurors who participated in the surveys are presented in Table 2. It can be seen that the demographics of the listening test participants strongly correlate with the desired demographics. The average age is close to that desired, and well over half of the jurors have at least intermediate sewing skills and sew on a rather regular basis. Due to constraints with the project, no detailed hearing tests of the jurors were able to be performed. Our assumption is that the hearing of the jurors was representative of the general population, and that some hearing loss was likely in at least some of the jurors.

3.2 Mathematical Sound Quality Metrics Assessment

The complexity of using humans as instruments for sound quality evaluation has motivated the development of equations that will calculate a set of metrics to describe the perceived sound quality of a sound. The metric chosen for investigation in this study is the sensory pleasantness, developed by Zwicker and Fastl[9] for other applications. This metric characterizes human perception of sound in terms of additional metrics, including loudness, sharpness, roughness, fluctuation and tonality. This metric will be outlined only briefly to specify the metrics that were used. For a more in-depth discussion of these metrics, the reader should refer to the more in depth discussion provided by Chatterley[10] and by Zwicker and Fastl[9].

Since the human auditory system exhibits a dependence on both amplitude and frequency, sound quality metrics are generally calculated using critical band rates (expressed in Bark) instead of frequency (in Hz). Critical bands are very similar to one-third octave bands above 500 Hz, but are nearly constant bandwidth below 500 Hz. The dependence of loudness on frequency is well-known, as shown by the equal loudness curves presented by Robinson and Dadson[11].

To calculate the loudness, the specific loudness (loudness per critical band) is first determined, as shown in Eqn. (1), after which the loudness is obtained as the integral of specific loudness, as shown in Eqn.(2).

$$N' = 0.08\left(\frac{E_{TQ}}{E_O}\right)^{0.23}\left[\left(0.5+\frac{E}{2\cdot E_{TQ}}\right)^{0.23}-1\right]\frac{sone}{Bark} \qquad (1)$$

$$N = \int_0^{24\,Bark} N'\,dz \qquad (2)$$

In Eqn. (1), E_{TQ} is the excitation level at the threshold of quiet, E_o is the excitation that corresponds to the reference intensity $I_o = 10^{-12}$ W/m^2 and E is the excitation level per critical band rate of the sound in question.

Sharpness is a sound quality metric used to quantify the high frequency content of the sound. It is computed as a weighted area of loudness, similar to an area moment calculation, as shown in Eqn. (3).

Table 2 - Jury demographics

No. of jurors	% female jurors	Average juror age	% jurors who know how to sew	% jurors who own a sewing machine	% jurors who sew monthly or more frequently	% jurors who have intermediate to advanced sewing skills
82	80.5%	35	91.5%	85.7%	54.9%	61.0%

$$S = 0.11\frac{\int_0^{24Bark} N' \cdot g(z) \cdot z dz}{\int_0^{24Bark} N' \; dz} \tag{3}$$

In Eqn. (3), $g(z)$ is a weighting function that has a unitary value below ~3 kHz and increases (non-linearly) from there to ~20 kHz, where it has a value of four. The variable z represents the critical band.

Roughness and fluctuation are similar metrics. They describe the modulation or audible oscillation of sounds. The roughness metric describes an auditory effect from rapid modulation frequencies, which coupled with the masking effects of the human auditory system, create a perception of grating or harshness to the sound. The dependence of roughness on frequency of modulation is greatest at a modulation frequency of 70 Hz, but the sensation begins in the range of 10 to 20 Hz and falls off from its maximum at 70 Hz to zero again between 400 and 500 Hz. The model for roughness is depicted in Eqn. (4).

$$R = 0.3\frac{f_{\mathrm{mod}}}{1kHz}\int_0^{24Bark}\frac{\Delta L_E(z)}{dB/Bark}dz \tag{4}$$

In Eqn. (4), f_{mod} is the modulation frequency, while $\Delta L_E(z)$ is related to the temporal masking depth. In their earlier work, Zwicker and Fastl[1] estimated this to be given by

$$\Delta L_E(z) = 20\log\left(\frac{N'_{\max}}{N'_{\min}}\right)dB \tag{5}$$

where $N'_{\max}$ and $N'_{\min}$ are the maximum and minimum specific loudness in the current critical band. This model then reduces to a sum of 24 terms representing the ratio of the maximum and minimum specific loudness in each critical band rate.

Similarly, fluctuation has a rising and falling relationship with frequency of modulation. However, it represents slower modulation frequencies. The perceptual effects of fluctuation are small for modulation frequencies near zero. These effects increase to a maximum at a modulation frequency of 4 Hz, and then drop off again to a minimal effect at modulation frequencies of about 25 to 30 Hz. The model of fluctuation strength is given in Eqn. (6),

$$F = \frac{0.008 \cdot \int_0^{24Bark}(\Delta L / dB\ Bark)dz}{\left(f_{\mathrm{mod}}/4Hz\right)+\left(4Hz/f_{\mathrm{mod}}\right)} \tag{6}$$

where f_{mod} is again the modulation frequency. Zwicker and Fastl[12] estimated the temporal masking depth, ΔL, to be given by

$$\Delta L = 4\log\left(N'_{\max}/N'_{\min}\right). \tag{7}$$

As is evident by the overlay of the two frequency bands (10Hz to ~400 Hz for roughness and 0 to ~30 Hz for fluctuation) there exists an area which is unclear as to which sensation it belongs to. Fluctuation frequencies are slow enough that the masking effects of the human auditory system only partially block the variation in amplitude. Often a sensation of motion is associated when a sound is strongly fluctuating, especially if the listener closes their eyes.

Tonality represents the absence or presence of strong tonal content in a wider band of sound. For this work, several methods for determining the tonal content were investigated[13,14] It was found that the tonality of the different sewing machines was nearly identical, and hence did not affect the ordering of the machines in the final sensory pleasantness metric that was used. As a result, tonality was not used as a parameter in our final model.

The sound quality metrics of loudness, sharpness, roughness, and fluctuation are then combined into a single "curve-fit" metric called sensory pleasantness[9]. This metric yields a single value representing the holistic acoustic appeal of the sound. Sensory pleasantness was originally derived from the other metrics by curve fitting the set of metrics to a myriad of juror test results. One of the objectives of this research was to determine if the overall sensory pleasantness metric is applicable for this significantly different application involving sewing machines. The benefit of this metric is that the relative importance of the individual metrics can be determined from the way a change in each metric affects the sensory pleasantness value. The model of sensory pleasantness is given in Eqn. (8).

$$\frac{P}{P_o} = e^{-\left(0.023\frac{N}{N_o}\right)^2} e^{-1.08\frac{S}{S_o}} e^{-0.7\frac{R}{R_o}}\left(1.24 - e^{-2.43\frac{T}{T_o}}\right) \tag{8}$$

In Eqn. (8), the subscript "*0*" represents the relative value that the sensory pleasantness of the sound under investigation is being compared to, P is the sensory pleasantness value, N the loudness, S the sharpness, R the roughness and T the tonality (which was set equal to T_0 for our analysis).

In summary, the sound quality metrics used in this research include loudness, sharpness, roughness, fluctuation strength, and sensory pleasantness. A very brief overview of these metrics has been presented for completeness; a complete derivation and justification for these models are provided by Zwicker and Fastl[9]. The sensory pleasantness results will be compared to the juror results to verify the correlation and validity of the juror tests.

3.3 Software Development and Verification

Using MatLab, sound quality metrics were calculated for each of the sewing conditions (both variation in speed and stitch pattern). Loudness was calculated using a set of MatLab m-file functions based on the ISO 532B/DIN45631 standard[15]. The remaining metrics (sharpness, roughness, fluctuation strength and tonality) were calculated using m-files created for this research. Tones and narrowband white noise of specific sound pressure level and center frequency, as well as bandwidth, were used to verify the accuracy of each of the m-files used to calculate the metrics.

Loudness verification was done using tones at several different sound pressure levels and different frequencies. The total loudness (in Sones) calculated from the tone was compared to the loudness level (Phon) for that frequency. Results for three of the test tones are listed in Table 3. The slight variation in the numbers is due to numerical round off.

Table 3 - Loudness verification

	100 Hz 50 dB	1 kHz 70 dB	2 kHz 60 dB
Equal loudness contours chart	1 sone 40 phon ~ 50 dB at 100 Hz	8 sone 70 phon ~ 70 dB at 1 kHz	5 sone 65 phone ~ 60 dB at 2 kHz
Calculated in MATLAB	1.04 sone	8.02 sone	4.97 sone

The MatLab m-files for sharpness, roughness and fluctuation were also coded using the theory described in Sec. 3.2. Verification tones, when not specifically stated in Ref. [9], were assumed to be root mean square (RMS) sound pressure levels.

Sharpness test tones of one critical bandwidth narrowband noises at 60 dB are used to evaluate the accuracy of the sharpness m-file. The results, compared to graphical data from Ref. [9], are tabulated in Table 4.

Roughness and fluctuation were verified using amplitude modulated tones with appropriate modulation frequencies. The results are compared to subjective test results in Ref. [9]. These results are displayed in Table 5, where they are compared to the calculated values from our MatLab code. As can be seen in the table, the results are a little off for both fluctuation and roughness. However, the general trend is correct and arguably within the error bounds of the subjective test results found on pages 252 and 258 of Zwicker and Fastl[9].

Table 4 - Sharpness code verification

Center Frequency	2 (Bark)	8.5 (Bark)	16 (Bark)	22.5 (Bark)
	200 Hz	1 kHz	3.15 kHz	10.5 kHz
Bandwidth	100 Hz	160 Hz	500 Hz	2.5 kHz
Standards	~0.25 (acum)	~1.0 (acum)	~2.0 (acum)	~8.0 (acum)
MATLAB calculated results	0.283 (acum)	0.996 (acum)	1.950 (acum)	6.498 (acum)

4 SOUND QUALITY RESULTS

This research evaluated the sound quality of six consumer sewing machines, from two brands. The machines ranged in sewing level from entry-level mechanical sewing machines to high-end computer controlled sewing machines. The nomenclature used in referring to the machines is outlined in Table 6.

4.1 Jury Listening Test Sound Quality Results

The repeatability of the juror responses determines the level of accuracy for each individual juror. This process allows the level of differences that the juror can recognize to be determined. Preliminary juror listening tests indicated that much of what was being examined would be difficult for the jurors to distinguish. Therefore, every comparison performed by each juror was repeated by that juror at least twice. This allowed the discovery of the "threshold of auditory discernability" for each comparison and alternatively each juror. It also allows an estimate of general consumer trends in a case such as this, where differences between machines are not clearly distinguished. Using these repeated questions, it was determined that on average the jurors were able to reproduce the same answer given the same sounds 60% of the time. This, as was suspected, is not a large percentage. It implies that the differences in the sounds are not substantial enough for the average listener to confidently select the same sound every time they hear the same comparison. These results imply that the jury listening test results may not lend a definitive result. However, there are general trends in the responses that are still valid and can be used to determine preferences. This is accomplished by ranking the percentages of preference for each machine irrespective of repeated questions. For each comparison comprising machine A and machine B, if the juror selects machine A as the preferred machine the first time a point is added to machine A's score. If the second time the two machines are compared, the same juror selects machine

Table 5 - Roughness & Fluctuation verification

	Fluctuation			Roughness		
	AM BBN 60 dB	AM SIN 70 dB	Narrow- Band Noise	AM SIN 125 Hz Center freq	AM SIN 1 kHz Center freq	AM SIN 8 kHz Center freq
Modulation Depth	80%	80%	N/A	100%	100%	100%
Modulation Frequency	4 Hz	4 Hz	N/A	40 Hz	70 Hz	100 Hz
Bandwidth	16 kHz	N/A	10 Hz	N/A	N/A	N/A
Standards	~1.50 Vacil	~1.25 Vacil	~0.25 Vacil	~0.35 Asper	1 Asper	~0.40 Asper

B, a point is added to machine B's score. At the end, the total number of comparisons is used to find the preference percentage. This results in a percentage-preferred ratio. Figure 3 is a bar graph of percentage-preferred results. This figure represents a summary of all the jury testing results.

As indicated in Fig. 3, the consumer trend appears to be towards a preference for the entry (E1 and E2) and mid level machine (M1) sounds, with the noted exception of the second mid level machine (M2). This is a strong contrast to the research hypothesis that the high-end machines would have the best sound quality, followed by the mid level and lastly the entry-level machines. Therefore, it was desired to also investigate mathematical sound quality metrics, in order to determine if these jury preferences could be reproduced with those metrics.

Additional evaluations performed by the jurors include that of the spectral reduction of the signal, as well as the addition of foam. Both of these are described in Table 1. The results for reduction in spectral content indicate that the consumer prefers a reduction in the higher frequencies. For these results, machine sounds with the high frequency content reduced were compared with machine sounds with the low frequency content reduced. It can be seen that there is a significant preference across the machines and speeds evaluated for the high frequencies to be attenuated, as is evident in Fig. 4.

The results for machine isolation from the table are not as clear. It is noted that the same table and mount position was used for each machine. The results indicate that the consumer preferences are dependant upon each machine (see Fig. 5). These results indicate that for three of the machines, the annoying sounds are not generated by the vibration transmission from the sewing machine to the table. For the other three models, H1, H2 and M2, vibrations transferred to the sewing table could be responsible for some of the unwanted noise. This is especially true for H1 and M2, where the consumer preferences were very strong.

The juror results indicate that our initial hypothesis is likely incorrect. In other words, the more acoustically pleasant machines are the entry-level machines, with the second brand entry-level machine being the most preferred, followed by the first brand entry-level machine. The jurors prefer the reduction of high frequency content to the reduction of low frequencies. Additionally, isolation is an issue for some of the machines, but is not a general issue for all six machines. Given these results, an analysis using mathematical metrics was carried out, to determine how those results compare with the jury results.

4.2 Mathematical Sound Quality Metric Results

The calculated sound quality metrics can be evaluated individually, when it is desired to focus on a single attribute

Table 6 - Machine nomenclature

Level	Brand	Acronym
High end	First	H1
	Second	H2
Mid-level	First	M1
	Second	M2
Entry-level	First	E1
	Second	E2

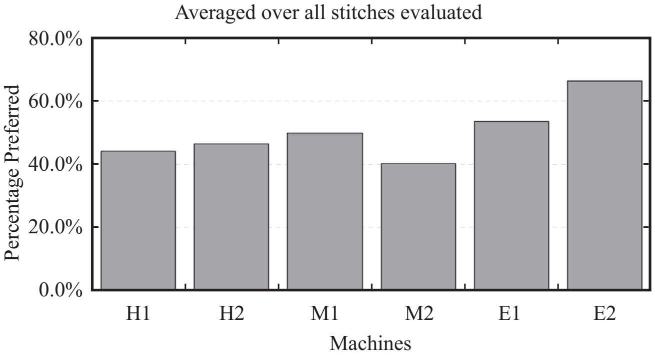


Fig. 3. Juror results comparing machines.

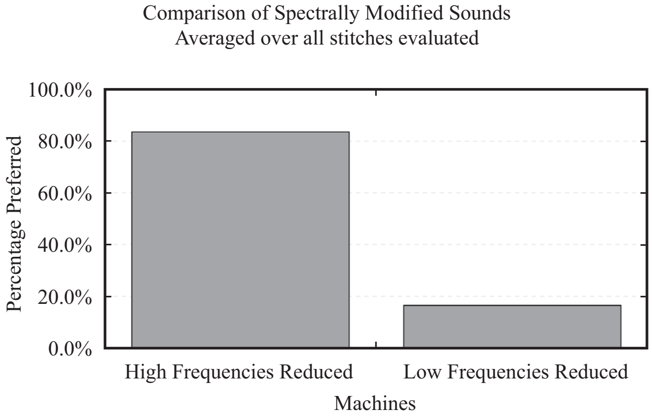


Fig. 4. Comparison of spectrally modified sounds.

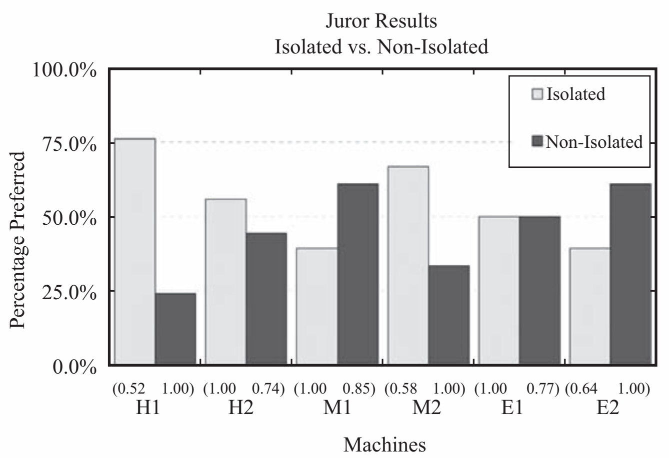


Fig. 5. Isolated with foam blocks vs. non-isolated sounds consumer preference percentage (light gray (left to right lines) represents the isolated sounds, black (right to left lines) non-isolated sounds). The metric based sensory pleasantness values for the corresponding machines are shown in parentheses below the bar chart values.

of the sound in question. Alternatively, they can be combined through the relative sensory pleasantness model. In the following results, all metric values have been normalized against the largest value. In each of the metrics, except sensory pleasantness, a lower number is desired. In sensory pleasantness, however, the larger numbers represent the acoustically pleasant sounds.

As previously mentioned, all of the metrics were calculated for each machine at each stitch and speed (straight, straight reinforced, zigzag, and zigzag reinforced at low, medium and high speeds). The numbers were averaged for each machine across the range of stitches. Table 7 presents the normalized results for all six of the machines. The final column in the table is the sensory pleasantness ranking for each machine. Sensory pleasantness was computed based on the values of the other four metrics to give a single overall rating of the sound quality for the given machine. To reiterate, in all of the metrics except sensory pleasantness, a lower number is desired. However, in sensory pleasantness, the larger value indicates a more appealing sound.

From Table 7, specifically the final column, it becomes evident that the most appealing sounds are coming from the low-end machines. It should also be noted that both tonality and fluctuation strength have almost no impact on the relative sensory pleasantness, as the variation from strongest to weakest in both categories was exceptionally small and did not yield differentiating information.

Additional calculated metric results include the evaluations of spectrally modified sounds and isolated vs. non-isolated machine sounds. The results of metric calculations for the spectrally modified sounds are presented in Table 8. Also included are the metric calculations averaged over all machines for the non-modified spectra. These metrics are an average over all the values presented in Table 7. From Table 8 it is evident that the jury results are again supported by the mathematical metrics, where a reduction in the high frequency content is preferred over a reduction in the low frequency content. In fact, even non-modified signals are preferred over

Table 7 - Sound quality metric results (Normalized)

Machine	Loudness	Sharpness	Roughness	Fluctuation	Tonality	Sensory Pleasant-ness	Rank
H1	0.86	1.0	0.90	1.0	0.98	0.19	5
H2	0.90	0.96	0.76	0.93	0.95	0.28	3
M1	0.78	0.82	0.99	1.0	0.99	0.21	4
M2	0.779	0.989	1.0	0.979	0.999	0.16	6
E1	1.0	0.72	0.44	0.99	1.0	0.78	2
E2	0.70	0.71	0.33	0.99	0.99	1.0	1

Table 8 - Sound quality metric results, with digital reduction of spectral content (Normalized)

	Reduced High Frequencies	No Change to the Frequency Spectrum	Reduced Low Frequencies
Loudness	0.92	0.94	1.0
Sharpness	0.88	1.0	1.0
Roughness	1.0	0.92	0.99
Fluctuation	0.92	0.91	1.0
Tonality	0.98	0.96	1.0
Sensory Pleasantness	1.0	0.96	0.80

a reduction in low frequency content. The implication from this is that it would be desirable to focus on modifications designed to reduce the high frequency sound production of the machine, as opposed to modifications that target primarily low frequency sound. The values in Table 8 are the average of all the sewing machines.

Another comparison performed by the jurors was the evaluation of sounds for a straight stitch medium speed where the machine was isolated from the table by the addition of foam rubber blocks under the feet. These sounds were compared to the standard machine sound (non-isolated). These sounds would approximate an improvement in the vibration isolation of the machine from the sewing table. The results for these tests are in Table 9.

Table 9 - Isolated vs. non-isolated sounds (Normalized)

	Isolated Sensory Pleasantness	Non –Isolated Sensory Pleasantness
H1	0.52	1.00
H2	1.00	0.74
M1	1.00	0.85
M2	0.58	1.00
E1	1.00	0.77
E2	0.64	1.00

In slight contrast to the juror results, it appears that H1 and M2 should not be more isolated and that M1 should be more isolated. This, however, is not surprising given that the jurors judged these sounds as very similar. Figure 6 shows a plot of the specific loudness for H1, where the dashed line represents the original sound (non–isolated) and the solid gray line represents the specific loudness with the addition of foam. It can be seen that the results with and without isolation visually appear to be rather similar. Thus, it is apparent that there are subtle differences that account for the differences in sensory pleasantness noted in Table 9. It should also be noted in Table 9 that the H1, H2, M1, and M2 machines do not compare with the metric calculations of sensory pleasantness but that the E1 and E2 machines do compare. This may be due to the amount of data that was acquired.

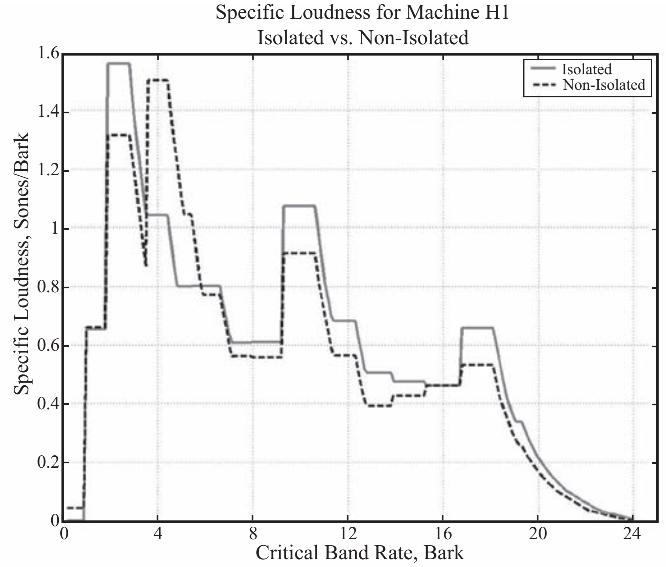


Fig. 6. Machine isolation specific loudness plot.

The final comparison is that of the start up sounds of each machine. The jurors were instructed to listen to the start up process, and when sewing commenced, they were to disregard the sewing sounds. The start up sounds evaluated in the metrics were truncated to exclude any actual sewing sounds and therefore represent just the machine initiation and registration sounds. The registration sounds occur on the high end machines, and are associated with the initialization of the computer controlled stepper motors. These stepper motors are different for each machine. The results are displayed in Table 10, in which the preferred sounds tend to be those without the sound registration (machines E2 and M1). However, a more in-depth approach would include variations to each of the machine start up procedures.

Table 10 - Machine start up results (Normalized)

Machine	Start Up Sensory Pleasantness
H1	0.66
H2	0.60
M1	0.83
M2	0.50
E1	0.62
E2	1.0

Although there are some minor variations in the results from jurors when compared to the calculated metrics, the general trend, especially when comparing the machines against one another, is supported by the metrics. Therefore, it must be concluded that the null hypothesis is incorrect and that the hierarchy of sound does not follow that of the machine. In other words, for the machines tested in this report, the sensory pleasantness of the high-end machines tested in this study is not perceived as being of a better quality than that of the entry-level machines tested.

5 CONCLUSIONS

Although sound quality is generally not specifically thought of by the consumer when they purchase a product, it does reflect on the consumer's perception of the overall quality of the product, its durability, strength and "solidness". The sound quality of six sewing machines, ranging from entry-level thru mid-level to high-end machines was evaluated using both juror listening tests and calculated metrics. The jury tests and the calculated metrics exhibit the same trends in most tests. It was found that the low-end machines are preferable to the high-end and mid-level machines. The results in this work illustrate that the jury listening tests and the calculated metrics both provide valuable information in product design.

From all of the jury surveys, measurements, and calculations, it can be inferred that the ideal sewing machine has certain attributes. First, it should be smooth sounding, yet have a rhythm. This can be interpreted to mean that there is relatively low roughness to the machine's sound, yet a high fluctuation appears to be acceptable. This is likely because the machine operating speeds all fall in the range of fluctuation. It should sound "solid," as many jurors noted on their questionnaires, suggesting that the sharpness of the machine should be relatively low. Many jurors also preferred the quieter machine. This obviously refers to the loudness of the machine, which should be "quiet enough that you can talk on the phone while using it," as one juror noted on her survey. Therefore, the direction for an optimal sound is to reduce the roughness and sharpness first, followed by the loudness and fluctuation. As the pool of sounds in this study is hardly reflective of all possible sewing machine sounds, it cannot be stated though that the best sounding machine in this group is ideal. More investigation with the addition of several other brands would allow a more precise definition of the ideal sound.

The results from the calculated metrics suggest that the roughness, sharpness, and loudness have the greatest impact on the overall sensory pleasantness, in that order. Changes that would reduce the roughness would have the greatest impact on sensory pleasantness, while reducing the high frequency content (sharpness) from the machine sound would also have a positive impact. Reducing the loudness would achieve the objective noted by the jurors of being able to talk on the phone while using the machine.

The results of this work have shown that the sound quality of the high-end sewing machines is lower than desired. To improve the sound quality of the high-end machines, it would be necessary to modify certain machine operations, processes or components to achieve the desired level of sound quality. Possible modifications include the improvement of the machine isolation from the table with improved spring-damper feet. The selection of alternate components, perhaps using involute chains and sprockets as opposed to cogged belts or gears is another possibility. Addition of more passive noise control techniques including vibration-isolating mounts for the machine skin could be examined. The machine operating system could be modified to ramp up stepper motors at initialization.

6 REFERENCES

[1] R.H. Lyon, *Designing for product sound quality* Marcel Dekker, (2000).

[2] J. Blauert, "Product-sound assessment: An enigmatic issue from the point of view of engineering," *Proc. Internoise 94*, pp. 857 –862, (1994).

[3] J. Blauert and U. Jekosch, "Sound-quality evaluation –A multi-layered problem," Acustica **83**, 747-753, (1997).

[4] R. Guski, "Psychological methods for evaluating sound-quality and assessing acoustic information," Acustica **83**, 765-774, (1997).

[5] J. Blauert and U. Jekosch, "A semiotic approach towards product sound quality," *Proc. Internoise 96*, pp. 2283 –2286, (1996).

[6] Noise Control Eng. J. **51**, (2003).

[7] R.H. Lyon, "Designing for sound-quality," *Proc. Internoise 94*, pp. 863–868, (1994).

[8] D. Bowen and R.H. Lyon, "Mapping perceptual attributes of sound to product design choices," Noise Control Eng. J. **51**, 271-279, (2003).

[9] E. Zwicker, H. Fastl, *Psychoacoustics: Facts and models*, 2nd Ed., Springer–Verlag, (1999).

[10] J. Chatterley, "Sound quality analysis of sewing machines," *MS Thesis, Brigham Young University*, (2005).

[11] D. W. Robinson and R. S. Dadson, "A re-determination of the equal-loudness relations for pure tones," *Br. J. Appl. Phys*. **7**, 166-181 (1956).

[12] E. Zwicker, and H. Fastl, *Psychoacoustics: Facts and models,* Springer–Verlag, (1990).

[13] A. Hastings and P. Davies, "An examination of Aures's model of tonality," *Proc. Internoise 02*, Paper N620 (6 pages on CD), (2002).

[14] T. L. Lagö, "Accurate amplitude measurements for combined tonal and random noise," *Proc. 6th International Symposium on Transportation Noise and Vibration*, (10 pages on CD), (2002).

[15] MatLab m-file functions based on the ISO 532B / DIN 45631 loudness calculation from BASIC Program Published in J. Acoust. Soc. Jpn (E) 12, 1 (1991) by E. Zwicker, H. Fastl, U. Widmann, K. Kurakata, S. Kuwano, and S. Namba. http://widget.ecn.purdue.edu/~hastinga/Research.htm