ARL-TR-7960 ● FEB 2016

US Army Research Laboratory

Two Classification Methods for Grouping Common Environmental Sounds in Terms of Perceived Pleasantness

by Kelly Dickerson, Brandon S Perelman, Laura Sherry, and Jeremy R Gaston

NOTICES

Disclaimers

ARL-TR-7960 ● FEB 2016

US Army Research Laboratory

Two Classification Methods for Grouping Common Environmental Sounds in Terms of Perceived Pleasantness

by Kelly Dickerson, Brandon S Perelman, and Jeremy R Gaston
Human Research and Engineering Directorate, ARL

Laura Sherry
Towson University, Towson, MD

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED (From - To)**
February 2016	Technical Report	November 2015–May 2016

4. TITLE AND SUBTITLE Two Classification Methods for Grouping Common Environmental Sounds in terms of Perceived Pleasantness	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Kelly Dickerson, Brandon S Perelman, Laura Sherry, and Jeremy R Gaston	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory ATTN: RDRL-HRF D Aberdeen Proving Ground, MD 21005-5425	**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-TR-7960
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Real-world sounds are perceived as more than just a collection of acoustic attributes. They contain both acoustic and semantic attributes, which together influence perception. Acoustic measurements are usually well-defined and broadly agreed upon, but semantic information can be difficult to operationalize. Semantic attributes contain information related to context and listener experience, attributes that are subjective and prone to individual differences. The current study had 2 goals: 1) characterization of a set of sounds in terms of their perceived pleasantness and 2) comparison of 2 classification methods, experimenter-defined and objectively defined (i.e., data-defined) classification. Fourteen listeners rated 36 common environmental sounds on a scale of 1 to 7 for perceived pleasantness. Overall, the set trended toward the unpleasant end of the ratings scale, and impulse sounds were rated as more pleasant than continuous sounds. Further, it appears that natural sounds were rated as more pleasant than man-made sounds, and this is consistent with both experimenter- and objectively defined classification approaches, suggesting a latent categorical structure for pleasantness.

15. SUBJECT TERMS

environmental sound perception, classification methods, categorization, subjective experiences, Likert scales

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Kelly Dickerson
					19b. TELEPHONE NUMBER (Include area code)
Unclassified	Unclassified	Unclassified	UU	24	410-278-5979

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures **iv**

List of Tables **iv**

1. Introduction **1**

1.1 Defining Environmental Sounds through Listener Experiences 2

1.2 Goal of the Present Study 3

2. Methods **5**

2.1 Participants 5

2.2 Materials and Stimuli 5

2.3 Procedure 6

2.4 Analysis Overview 6

3. Results and Discussion **7**

3.1 Descriptive Statistics 7

3.2 Experimenter-Defined Categories 7

3.3 Cluster-Defined Categories 8

4. Conclusions **11**

5. References **13**

List of Symbols, Abbreviations, and Acronyms **16**

Distribution List **17**

List of Figures

Fig. 1 Mean pleasantness ratings for the 4 experimenter-defined categories. The difference in pleasantness between mechanical and natural sounds was significant, as was the difference in pleasantness between continuous and impulse sounds. 8

Fig. 2 Two-dimensional MDS solution of this stimulus set. Color assignments indicate the 2 emergent clusters obtained using stepwise FMM. The cluster indicated in black contained mainly mechanical sounds, while the cluster indicated in red contained bells, animal sounds, instrument sounds, and water movement sounds. The mean pleasantness ratings for sounds in the black cluster was 2.49 (*SE* = 0.07), and in the red cluster, which was primarily natural sound, the mean pleasantness rating was 4.73 (*SE* = 0.13)....9

Fig. 3 BIC for univariate mixture models containing different numbers of components of differing variance. The lowest BIC value indicates the mixture of models that provides the best fit, penalized by the number of components. 10

List of Tables

Table 1 Means and standard error for each of the 36 sounds 4

Table 2 Means and standard errors for natural and mechanical sounds 6

Table 3 Means and standard errors for continuous and impulse sounds 6

1. Introduction

Much of the hearing research for the US Army has been related to noise exposure and mitigating noise hearing hazard (see Scharine et al. 2005 and Fedele et al. 2013 for examples). However, the operational environment for Soldiers is more than just loud; operational environments contain a complex and dynamic auditory milieu. Common everyday environmental sounds may be considered just noise in the background but could potentially provide Soldiers with important operational information and enhance situational awareness. For example, changes in traffic patterns or cell phone activity can indicate danger. This acoustic milieu is ever-present in the urban environments in which Soldiers find themselves and represents a dynamic and continuously changing operational context that is currently not well understood. Environmental sounds can be defined as nonspeech, nonmusical signals that are meaningful as they relate to objects or events in the environment (Ballas and Howard 1987; Giordano et al. 2010). These environmental sounds are often acoustically complex, with dynamically varying temporal and spectral properties that are defined by the mechanics of the sound-producing event (see Fletcher and Rossing 1998 and Peirce et al. 1998 for examples).

This complexity is compounded by the fact that environmental sounds in naturalistic settings are rarely heard in isolation. This distinction is important because performance on localization, identification, and detection tasks is poorer when a sound is presented in the context of other auditory stimuli (i.e., an auditory scene) than when it is presented in isolation (for a detailed review of the influence of auditory context on performance, see Yost and Fey 2007 or Yost 2008). Specifically, the increase in complexity in the auditory background can contribute to overall perceptual load and influence performance on a variety of other auditory perceptual tasks (see Dickerson and Gaston 2014 for review) such that as perceptual load increases, performance deteriorates. For example, Leech et al. (2009) found that listener identification accuracy for 23 common environmental sounds was better when the sounds were presented in isolation than when presented in a competing background. Further, the content of the background also influenced performance. Listeners' accuracy improved when the to-be-identified sound was distinct from the background as opposed to when it was congruent or similar to the background (distinctiveness was defined in terms of similarity, which was evaluated separately in Gygi and Shafiro 2007). Further, semantic-level effects do not appear to be limited to identification performance. Recently Dickerson et al. (2016) found that localization and change discrimination performance were poorer when the target was similar to the background than when the target was distinct (dissimilar) from the background (see also Gregg and Samuel 2009 for related finding).

The previous examples indicate that real-world auditory perception is a complex process that cannot be captured by isolating single stimulus dimensions in the laboratory. Measurement of acoustic dimensions such as frequency and amplitude, and their perceptual correlates pitch and loudness, do not always account for differences in change discrimination performance for environmental sounds. Gregg and Samuel (2009) found that both semantic and acoustic information influences change discrimination performance, but that for meaningful environmental sounds the semantic information appears to account for listener performance better than acoustic information alone. Their finding suggests that in the absences of semantic information, listeners will rely on acoustic features, but that the default may be higher-order semantic features. Despite this, the fact remains that semantic features are more difficult to operationally define because of the likely influence of subjective factors such as similarity, identifiability, familiarity, and pleasantness. For example, when standing on a street corner, the experience includes things such as car doors slamming, traffic, birds chirping, and footsteps and not a series of low- and high-frequency intermittent impulses. That is, everyday auditory signals are perceived as object-oriented events. This object- or event-oriented listening means that acoustic information is blended with semantic and linguistic information, making stimulus characterization and classification tasks much more complex.

1.1 Defining Environmental Sounds through Listener Experiences

As was highlighted in the previous section, environmental sounds are different in terms of complexity and informational content compared with traditional laboratory stimuli. Much of the previous research on environmental sounds has been focused on developing taxonomies that map out the relationships among sounds in an attempt to uncover some underlying feature space (see Gaver 1993; Marcell et al. 2000; Gygi et al. 2007; Lemaitre et al. 2010; and Misdariis et al. 2010, for examples). While these taxonomies have been useful in mapping out the physical parameters of the sound-producing event, they do not necessarily reflect how a listener would group sounds based on their subjective experiences with a particular sound or cluster them within these hierarchical taxonomies. An alternative approach that addresses this limitation is to change focus from the experimenter-derived clustering taxonomies toward a more objective classification of listener perceptual experiences.

Stimulus selection decisions along with the characterization of the relationships between items within a stimulus set are often determined by a single experimenter with only the goals of the present study in mind. Fiedler (2011), in a review describing some common pitfalls in traditional psychological research, points out that experimenters, like their participants, are subject to context effects and bias. Fiedler suggests that the hallmark of "intuitive stimulus selection" may actually produce skewed and nonrepresentative stimulus sets. To begin to address this concern for an existing stimulus set, the present study examines differences in stimulus grouping based on 2 common methods: 1) a priori experimenter-defined groups and 2) data-driven clustering accomplished by applying a clustering algorithm, in this case Flexible Mixture Modeling (FMM), over a spatial mapping produced by Multidimensional Scaling (MDS). This approach of obtaining classification using clustering algorithms over MDS solutions is common in the auditory perception literature (Cermak and Cornillon 1976, Howard 1977, Gygi et al. 2007; Chang et al. 2010). If intuitive selection, that is, experimenter-defined stimulus groups, is representative of the groupings produced by participant observations, there should be no qualitative differences between the experimenter defined groups and those obtained using MDS and FMM.

1.2 Goal of the Present Study

Pleasantness is highly correlated with familiarity (Marcell et al. 2000; Bonebright 2001; Hocking et al. 2013), and this is the primary reason for its selection for study. Familiarity effects (e.g., perceptual and conceptual fluency, false memories, and other fluency heuristics) can profoundly influence judgments at multiple levels, from low-level perceptual performance such as detection or identification to higher-level decisions and assessments of risk. A pleasant-sounding signal may bias a listener toward an assessment of familiarity and lead to an inaccurate assessment of risk in a given situation. Thus, it is important to understand how familiarity operates directly by measuring familiarity via fluency in recall manipulations but also, indirectly, by measuring pleasantness.

The present study had 2 goals. The first goal was to map out the pleasantness space for a set of 36 common environmental sounds (Table 1). These 36 sounds fall into several distinct semantic categories and broadly represent an outdoor urban environment, a space where Soldiers often operate. By capturing basic normative data, such as pleasantness, it is possible to evaluate perceptual performance for these sounds in the context of their subjective attributes. The pleasantness data obtained in this normative study is part of a larger set of studies tying several types of subjective stimulus attributes to performance on perceptual and memory-related tasks.

Table 1 Means and standard error (SE) for each of the 36 sounds

Stimulus	Mean	SE
AlarmClock	1.50	0.091
Baby1	3.16	0.058
Baby2	3.14	0.070
Bell1	3.33	0.077
Bell2	3.47	0.061
Bike1	3.11	0.065
Bike2	2.69	0.043
Bus1	1.79	0.029
Bus2	1.93	0.043
Cans1	2.90	0.067
Cell1	2.86	0.064
Crickets1	4.10	0.063
Crickets2	4.90	0.047
Dog1	2.94	0.060
Dog3	4.26	0.055
DuckCall	3.01	0.032
Guitar	6.63	0.036
Helicopter1	2.54	0.038
Helicopter2	2.36	0.062
Jackhammer1	1.99	0.072
Lighter1	4.49	0.043
Metal1	2.71	0.066
Motorcycle2	2.86	0.055
Plane1	2.69	0.073
Plane3	2.69	0.042
Pouring1	5.72	0.075
Rain	6.47	0.102
Shopvac1	1.76	0.036
Shopvac2	1.96	0.070
Stream	6.47	0.083
Tank1	2.13	0.049
TeaKettle	3.99	0.063
Truck1	2.16	0.059
Truck2	2.29	0.048
Walking1	3.73	0.057
Walking2	3.95	0.058

In addition to the broader goal of characterizing the sound set in terms of pleasantness, 2 specific hypotheses were tested. There is some suggestion that despite the fact that most listeners have more experience with mechanical and man-made sounds, listeners actually prefer to hear natural sounds such as rainfall, animal calls, and footsteps over mechanical sounds (Marcell et al. 2000). The current study compares the pleasantness ratings of sounds a priori classified as natural and mechanical to test this hypothesis. The second hypothesis, that continuous sounds would be rated as more pleasant than intermittent or impulse sounds, is also based on previous studies suggesting that continuous sounds are preferred to intermittent and impulse sounds (see Hocking et al. 2013 for example) and will be evaluated using the same methods described for the first hypothesis.

The second goal of this study was to compare experimenter- and clustering-defined stimulus classification. The clustering algorithms should directly reflect the overall classification structure held by the participants, while the experimenter-defined clusters are potentially influenced by bias or experiences that may not be consistent with the participant ratings. By comparing experimenter- and clustering-defined groupings consistency in terms of coherence and semantic congruency among category, members can be compared across the 2 approaches.

2. Methods

2.1 Participants

Fourteen undergraduate students from the State University of New York at Binghamton participated in this study for course credit. All participants were given a description of the study and provided informed consent before beginning the study. After providing consent, participants' hearing was screened. Pure tone air conduction thresholds of 25 dB HL (hearing level) or better were measured in all participants at all octave frequencies between 500 and 8000 Hz prior to beginning the experiment to ensure normal hearing sensitivity.

2.2 Materials and Stimuli

Testing was conducted in a quiet room. Participants were seated at a laptop computer running E-Prime, and stimuli were presented at 70 dBA (A-weighted decibels) over Beyerdynamics DT 770 headphones.

Thirty-six sound stimuli were used. Table 1 describes the set of environmental sound stimuli. Dickerson et al. (2016) previously used 18 of the 36 stimuli. An additional 18 stimuli were collected from the website www.freesound.org. The sounds used in this study were selected because they were generally representative of an outdoor urban environment, and the range of sound sources were selected to encourage participants to use the entire rating scale. Animal vocalizations and natural sounds as well as sounds made by vehicles and construction tools were all included (Table 1). Out of the 36 sounds, 8 were used for the comparison between natural and mechanical (Table 2), and 14 were used in the comparison between continuous and impulse sounds (Table 3) described in the Results section.

Table 2 Means and standard errors for natural and mechanical sounds

Mechanical			Natural		
Stimulus	**Mean**	**SE**	**Stimulus**	**Mean**	**SE**
Lighter1.wav	4.49	0.04	Rain.wav	6.47	0.10
Jackhammer1.wav	1.99	0.07	Stream.wav	6.47	0.08
Bike1.wav	3.11	0.07	Crickets1.wav	4.10	0.06
Bike2.wav	2.69	0.04	Crickets2.wav	4.90	0.04

Table 3 Means and standard errors for continuous and impulse sounds

Continuous			Impulse		
SoundFiles	**Mean**	**SE**	**SoundFiles**	**Mean**	**SE**
Jackhammer1	1.99	0.07	Bell1	3.33	0.08
Tank1	2.13	0.05	Bell2	3.47	0.06
Shopvac1	1.76	0.04	Bike1	3.11	0.07
Shopvac2	1.96	0.07	Bike2	2.69	0.04
Truck1	2.16	0.06	Baby1	3.16	0.06
Bus1	1.79	0.03	Baby2	3.14	0.07
Bus2	1.93	0.04	Dog1	2.94	0.06

All sounds were truncated to 1000 ms in duration, with 5-ms linear on and off ramps to minimize acoustic transients. The entire set of sounds was normalized for root mean square amplitude in an effort to minimize amplitude differences across the set of sounds. All sound modifications were performed using Adobe Audition (CS 6).

2.3 Procedure

Following hearing screening and informed consent, listeners were seated in front of a laptop computer and donned the Beyerdynamic headphones to begin the experiment. On each trial a single sound from the set of 36 sounds was played. Listeners were then prompted to rate the sound using a 7-point Likert-type rating scale using the computer keyboard. A rating of 1 indicated that the sound was very unpleasant and 7 indicated that the sound was very pleasant. Each of the 36 sounds (Table 1) was randomly presented 5 times for a total of 180 trials.

2.4 Analysis Overview

To compare pleasantness ratings between experimenter- and data-defined groupings, we calculated descriptive statistics for the pleasantness ratings, then used inferential statistics to compare between experimenter-defined groups. Objective, data-defined classifications were obtained by the following process. First, we computed a 36*36 dissimilarity matrix using average pairwise differences in pleasantness for each pairwise comparison. Then, Kruskal's nonmetric MDS algorithm (Kruskal 1964) was used to map the 36 stimuli in a 2-D space. Finally, FMM provided a means of clustering the stimuli in the MDS solution (Leisch 2004; Gruen and Leisch 2007). FMM searches for emergent clusters within a solution by iteratively fitting a fixed number of Gaussian models to the data using the expectation-maximization algorithm, which rewards goodness of fit but penalizes the number of models required to fit the data. The number of models, *k*, can either be user-specified or inferred in a stepwise fashion by specifying a range for *k* and then taking the best *k* value according to metrics such as the Akaike Information Criterion or Bayes Information Criterion (BIC). The current study uses BIC, which enables data-driven cluster determination rather than experimenter interpretation of silhouette plot as is standard for other algorithms such as k-means (Lloyd 1982). The shape, orientation, and uniformity of these models can be modified to account for different data distributions.

3. Results and Discussion

3.1 Descriptive Statistics

Across the entire set of sounds, the average rating trended toward slightly unpleasant (M = 3.30, SE = 0.22); however, the range was quite broad, with the lowest pleasantness rating for alarm clock (M = 1.50, SE = 0.091) and the highest rating for guitar (M = 6.63, SE = 0.036). See Table 1 for mean pleasantness ratings for all sounds.

3.2 Experimenter-Defined Categories

Recall that the present study was interested in comparing experimenter-defined categories with those categories that emerged from the MDS analysis. The following subsections present the results from the a priori experimenter-defined comparisons between natural and man-made sounds and between continuous and impulse sounds.

Were natural sounds rated to be more pleasant than man-made sounds? The 36 sounds selected for inclusion in the rating task represented a broad array of everyday environmental sounds. Two semantic categories were identified as potentially different from one another: natural sounds and mechanical sounds (see Table 1 for means). A paired samples t-test revealed a large and significant effect of pleasantness on the differences between the 2 groups of sounds $t(3) = -3.27$, $p < 0.05$, (Cohen's d = 1.48, M_{diff} = 2.41) with natural sounds rated as significantly more pleasant (M = 5.49, SE = 0.59) than mechanical sounds (M = 3.07, SE = 0.52) These results are depicted in the left side of Fig. 1.

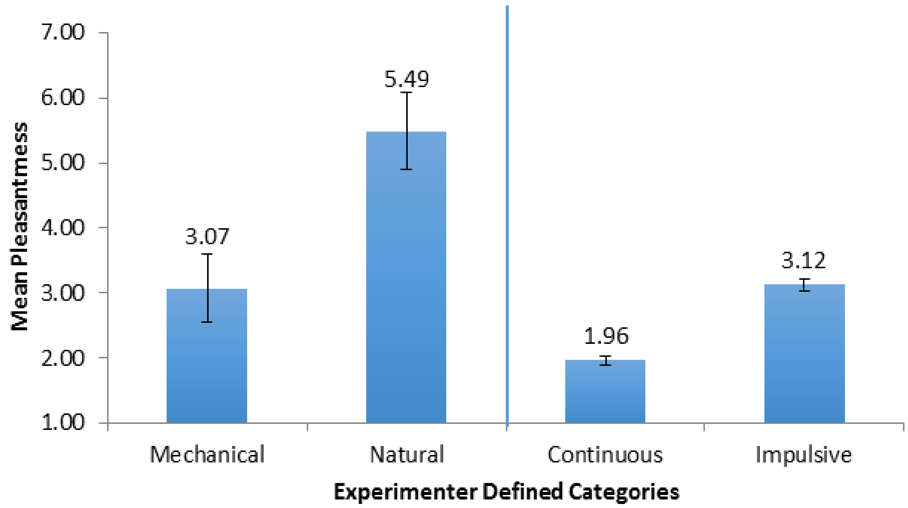

Fig. 1 Mean pleasantness ratings for the 4 experimenter-defined categories. The difference in pleasantness between mechanical and natural sounds was significant, as was the difference in pleasantness between continuous and impulse sounds.

Were continuous sounds rated to be more pleasant than impulse sounds? From the set of 36, 7 sounds were identified as impulse and another 7 as continuous. A paired samples t-test revealed a large and significant effect of pleasantness between the 2 groups, $t(6) = -10.19$, $p < 0.01$, (Cohen's $d = 0.94$, $M_{diff} = -1.19$). Impulse sounds were rated as significantly more pleasant ($M = 3.12$, $SE = 0.10$) than continuous sounds ($M = 1.96$, $SE = 0.06$). This effect is depicted in left side of Fig. 1. It is important to note that the sounds selected for this comparison are all man-made, thus the conclusions drawn about the pleasantness of impulse versus continuous sounds should be interpreted cautiously, as it may not extend to other sound categories, such as natural sounds.

3.3 Cluster-Defined Categories

Analysis of experimenter-defined categories supports the hypothesis that natural sounds are perceived as more pleasant than those that are man-made. The hypothesis that continuous sounds would be rated as more pleasant than impulse sounds was not supported by the data. This may have occurred, in part, because the analysis only included sounds that were defined a priori to fall into those particular categories and does not reflect the continuum of values along those (or other latent) dimensions. The next set of analyses takes the entire set of 36 sounds into account and clusters them based on commonality along 2 latent dimensions that contribute to the overall perception of pleasantness.

Nonmetric MDS space using a matrix composed of pairwise difference scores produced an MDS solution of minimal stress (2-D Stress = 0.01; permitting additional dimensionality did not reduce this value), allowing us to visualize the data in 2 dimensions (Fig. 2). Examination of the MDS solution reveals that the majority of the variance lies on the X axis. Therefore, we clustered the data using univariate stepwise FMM on the X axis coordinates only. We calculated BIC for 5 mixtures of models, containing 1–5 clusters; variance between clusters was permitted to be unequal. The 2-cluster model provided the best fit as measured by BIC (Fig. 3).

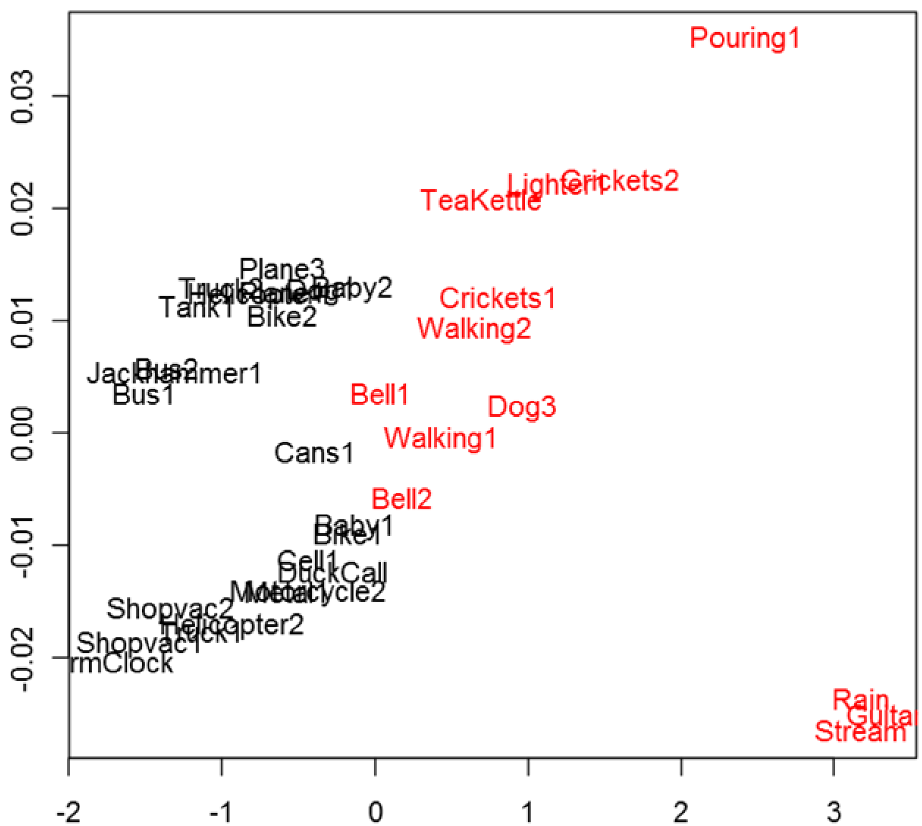

Fig. 2 Two-dimensional MDS solution of this stimulus set. Color assignments indicate the 2 emergent clusters obtained using stepwise FMM. The cluster indicated in black contained mainly mechanical sounds, while the cluster indicated in red contained bells, animal sounds, instrument sounds, and water movement sounds. The mean pleasantness ratings for sounds in the black cluster was 2.49 (*SE* = 0.07), and in the red cluster, which was primarily natural sound, the mean pleasantness rating was 4.73 (*SE* = 0.13)

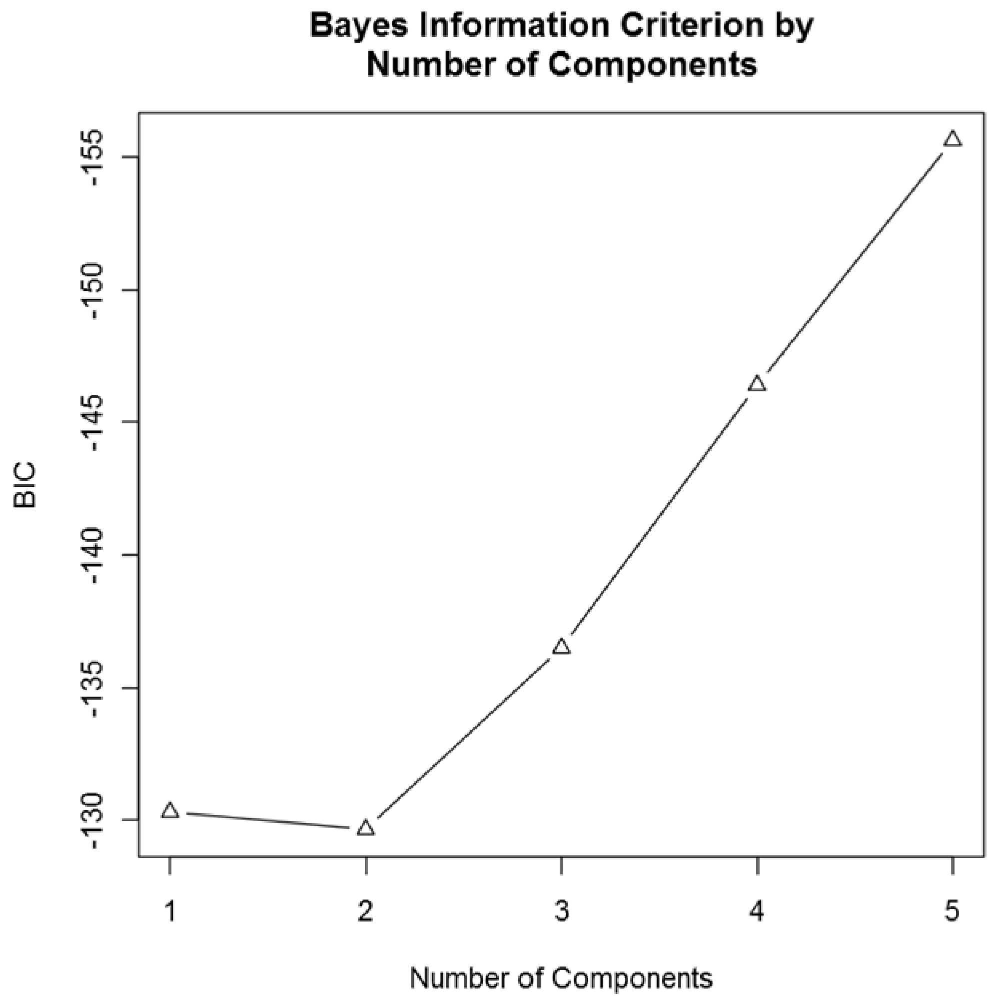

Fig. 3 BIC for univariate mixture models containing different numbers of components of differing variance. The lowest BIC value indicates the mixture of models that provides the best fit, penalized by the number of components.

The first cluster, depicted in black in Fig. 2, contained mainly mechanical sounds (n = 23/36), while the second cluster, depicted in red in Fig. 3 (n = 13/36), contained a mix of sounds made by instruments, animals, and water movement. The second cluster was rated as far more pleasant (M = 4.73, SE = 0.33) than the first (M = 2.49, SE = 0.10). A t-test was used to compare the ratings distributions for items belonging to each of the emergent clusters. The results of this t-test were significant [$t(295.24) = 15.51$, $p < 0.001$] (Cohen's d = 2.65, M_{diff} = 2.08), indicating that the emergent clusters identified by the mixture model adequately described differences in pleasantness between the 2 groups of stimuli. Overall, these results and those from the experimenter-defined classification suggest that stimulus pleasantness can be captured at a categorical level but that the specifics of category membership and structure depend on the classification methods selected.

4. Conclusions

The pleasantness of an environmental sound is an important stimulus attribute that likely contributes to situation awareness by influencing performance on auditory and other perceptual tasks. The present study revealed that natural sounds were perceived as more pleasant than man-made sounds and that impulse sounds were perceived as more pleasant than continuous sounds. However, the influence of temporal aspects of the stimuli on perceived pleasantness is difficult to evaluate because a significant proportion of the continuous sounds were also man-made sounds, which are rated as less pleasant in general. The contribution of pleasantness to situation awareness may be driven by its link to stimulus familiarity (Wagner and Gabrieli 1998; Marcell et al. 2000; Dickerson et al. in prep), which has been shown to bias performance, particularly on memory tasks. For example, familiarity biases memory recall performance such that participants falsely recollect familiar items a greater proportion of the time than unfamiliar items (Verde et al. 2007). These potential biasing effects of familiarity and pleasantness may detrimentally impact situation awareness by affecting Soldiers' ability to accurately evaluate the content and threat of an auditory scene. For example, salience can be impacted by the emotional valiance associated with a signal or event. Soldiers recall ability could be similarly impacted by pleasantness and familiarity. Thus, understanding the link between pleasantness and familiarity and its influence on the performance of other tasks is an important future direction for this work.

The present study evaluated the pleasantness of 36 common environmental sounds. These sounds were classified based on a priori experimenter-defined hypothetical categories and an objective clustering method that uses FMM to group stimuli based on their similarity relationships (i.e., distance) in an MDS space. Visual inspection of the differences between experimenter- and clustering-defined categories revealed some similarity in category membership between the 2 classification methods, but the categories differed in terms of their specific structure and composition. FMM, by default, used all available stimuli to create the 2-cluster solution illustrated in Fig. 2. The experimenter-defined categories were different because they were based on specific hypotheses about subtypes of sounds (continuous vs. impulse and natural vs. mechanical), thus the set of sounds used to create categories was down-selected from the set of 36 based on the experimenter's own internal criteria for items that would fit those categories.

While this type of stimulus selection decision is common, it is fraught with bias. It is the potential for biased stimulus selection that led to the explicit comparison between the subjective experimenter-defined classification and the more objective MDS-based clustering approach. This concern about biased stimulus selection influencing study results is not new. In a recent review, Fiedler (2011) describes the potential for statistics that are inflated or results that do not actually represent the distribution of variables in the environment due to experimenter biases in stimulus selection or study design. “Intuitive selection” of stimulus is lauded as a virtue, but paired with pilot testing can lead to problematically optimized stimulus sets that are designed to elicit an effect but do not represent human performance or the natural environment. By comparing experimenter- and clustering-defined category structures, the present study reveals that, per Fiedler’s, language, “intuitive” stimulus selection, at least in terms of the pleasantness of a stimulus, is not particularly biased. The classifications derived by experimenter-based selection were well represented in the data-derived FMM clustering.

The present study demonstrates that there are differences in perceived pleasantness between sounds classified as natural versus man-made. Further research is needed to determine how much specific stimulus features, whether acoustic, semantic, or both, drive the perceived pleasantness of common environmental sounds. Understanding the link between subjective stimulus evaluations and perceived pleasantness on perceptual performance has important implications for Soldier situation awareness. Soldiers are increasingly operating in dynamic and acoustically rich urban environments. It is the goal of this research and other studies like this to determine the extent to which background auditory information can be used to uncover meaningful and situationally relevant changes in the operational environment.

5. References

Ballas JA, Howard JH. Interpreting the language of environmental sounds. Environment and Behavior. 1987;19(1):91–114.

Bonebright TL. Perceptual structure of everyday sounds: a multidimensional scaling approach. Proceedings of the 2001 International Conference on Auditory Display. 2001;1:73–78.

Cermak GW, Cornillon PC. Multidimensional analyses of judgments about traffic noise. The Journal of the Acoustical Society of America. 1976;59:1412–1420.

Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. Nature Neuroscience. 2010;13(11):1428–1432.

Dickerson K, Gaston J R. Did you hear that? The role of stimulus similarity and uncertainty in auditory change deafness. Frontiers in Psychology. 2014:5:1125.

Dickerson K, Gaston JR, Perelman BS, Mermagen T, Foots AN. Sound source similarity influences change perception in complex scenes. In: Proceedings of Meetings on Acoustics – Acoustical Society of America. 2016;23:5–11.

Dickerson K, Sherry L, Nagel S. Eliciting the fluency heuristic using pleasantness as a proxy variable for familiarity: examining pleasantness and recall for common environmental sounds. Proceedings of Meetings on Acoustics – Acoustical Society of America. In prep.

Fedele PD, Binseel MS, Kalb JT, Price GR. MIL–STD–1474E: using the auditory hazard assessment algorithm for humans (AHAAH) with hearing protection software. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2013.

Fiedler K. Voodoo correlations are everywhere – not only in neuroscience. Perspectives on Psychology. 2011;6(2):163–171.

Fletcher NH, Rossing T. The physics of musical instruments. New York (NY): Springer; 1998.

Gaver WW. What in the world do we hear? An ecological approach to auditory event perception. Ecological Psychology. 1993;5(1):1–29.

Giordano BL, McDonnell J, McAdams S. Hearing living symbols and nonliving icons: category specificities in the cognitive processing of environmental sounds. Brain and Cognition. 2010;73(1):7–19.

Gregg MK, Samuel AG. The importance of semantics in auditory representations. Attention, Perception, & Psychophysics. 2009;71(3):607–619.

Gruen B, Leisch F. Fitting finite mixtures of generalized linear regressions in R. Computational Statistics & Data Analysis. 2007;51(11):5247–5252.

Gygi B, Shafiro V. Effect of auditory context on the identification of environmental sounds. Proceedings of the 19th International Congress on Acoustics; 2007 Sep 2–7; Madrid, Spain. p. 96, 98, 118.

Gygi B, Kidd GR, Watson CS. Similarity and categorization of environmental sounds. Perception & Psychophysics. 2007;69(6):839–855.

Hocking J, Dzafic I, Kazovsky M, Copland DA. NESSTI: norms for environmental sound stimuli. PloS One. 2013;8(9):e73382.

Howard JH Jr. Psychophysical structure of eight complex underwater sounds. The Journal of the Acoustical Society of America. 1977;62:149–156.

Kruskal JB. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrika. 1964;29(1):1–27.

Lemaitre G, Houix O, Misdariis N, Susini P. Listener expertise and sound identification influence the categorization of environmental sounds. Journal of Experimental Psychology: Applied. 2010;16(1):6.

Leech R, Gygi B, Aydelott J, Dick F. Informational factors in identifying environmental sounds in natural auditory scenes. Journal of the Acoustical Society of America. 2009;126(6):3147–3155.

Leisch F. FlexMix: A general framework for finite mixture models and latent class regression in R. Journal of Statistical Software. 2004;11(8):1–18.

Lloyd S. Least squares quantization in PCM. IEEE Transactions on Information Theory. 1982;28(2):129–137.

Marcell MM, Borella D, Greene M, Kerr E, Rogers S. Confrontation naming of environmental sounds. Journal of Clinical and Experimental Neuropsychology. 2000;22(6):830–864.

Misdariis N, Minard A, Susini P, Lemaitre G, McAdams S, Parizet E. Environmental sound perception: metadescription and modeling based on independent primary studies. EURASIP Journal on Audio, Speech, and Music Processing. 2010;(1):1.

Peirce CS, Houser N, Kloesel CJW. The essential Peirece: selected philosophical writings. Vol. 2. Bloomington (IN): Indiana University Press; 1998. p. 1893–1913.

Scharine AA, Henry PP, Binseel MS. An evaluation of selected communications assemblies and hearing protection systems: a field study conducted for the future force warrior integrated headgear integrated process team. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2005. Report No.: ARL-TR-3475.

Verde MF, Rotello CM. Memory strength and the decision process in recognition memory. Memory & Cognition. 2007;35(2):254–262.

Wagner AD, Gabrieli JD. (1998). On the relationship between recognition familiarity and perceptual fluency: evidence for distinct mnemonic processes. Acta Psychologica. 1998;98(2):211–230.

Yost WA, Fay RR. Auditory perception of sound sources. Vol. 29. Berlin (Germany): Springer Science & Business Media; 2007.

Yost WA. Auditory localization and scene perception. In: Goldstein EB, editor. The Blackwell handbook of sensation and perception. Hoboken (NJ): Wiley-Blackwell; 2008. p. 437.

List of Symbols, Abbreviations, and Acronyms

2-D	2-dimensional
BIC	Bayes Information Criterion
dBA	A-weighted decibels
FMM	Flexible Mixture Modeling
HL	hearing level
MDS	Multidimensional Scaling
SE	standard error

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

2 (PDF) DIRECTOR
US ARMY RESEARCH LAB
RDRL CIO L
IMAL HRA MAIL & RECORDS MGMT

1 (PDF) GOVT PRINTG OFC
A MALHOTRA

1 (PDF) ARMY RSCH LAB – HRED
RDRL HRB B
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL 35898-7290

7 (PDF) ARMY RSCH LAB – HRED
SFC PAUL RAY SMITH CENTER
RDRL HRO COL H BUHL
RDRL HRA I MARTINEZ
RDRL HRR R SOTTILARE
RDRL HRA C RODRIGUEZ
RDRL HRA B J HART
RDRL HRA A C METEVIER
RDRL HRA D B PETTIT
12423 RESEARCH PARKWAY
ORLANDO FL 32826

1 (PDF) USA ARMY G1
DAPE HSI B KNAPP
300 ARMY PENTAGON
RM 2C489
WASHINGTON DC 20310-0300

1 (PDF) USAF 711 HPW
711 HPW/RH K GEISS
2698 G ST BLDG 190
WRIGHT PATTERSON AFB OH 45433-7604

1 (PDF) USN ONR
ONR CODE 34 P MASON
875 N RANDOLPH STREET
SUITE 1425
ARLINGTON VA 22203-1995

1 (PDF) USN ONR
ONR CODE 341 J TANGNEY
875 N RANDOLPH STREET
BLDG 87
ARLINGTON VA 22203-1986

1 (PDF) USA HQDA
ARI M SAMS
6000 6TH STREET BLDG 1464
FT BELVOIR VA 22060

1 (PDF) USA NSRDEC
RDNS D D TAMILIO
10 GENERAL GREENE AVE
NATICK MA 01760-2642

1 (PDF) OSD OUSD ATL
HPT&B B PETRO
4800 MARK CENTER DRIVE
SUITE 17E08
ALEXANDRIA VA 22350

ABERDEEN PROVING GROUND

12 (PDF) DIR USARL
RDRL HR
L ALLENDER
P FRANASZCZUK
RDRL HRB
J LOCKETT
RDRL HRB A
M LAFIANDRA
RDRL HRB C
J GRYNOVICKI
RDRL HRB D
D HEADLEY
RDRL HRF
K OIE
RDRL HRF A
A DECOSTANZA
RDRL HRF B
A EVANS
RDRL HRF C
J GASTON
RDRL HRF D
A MARATHE
K DICKERSON

INTENTIONALLY LEFT BLANK.