Automating Ultrasonic Vocalization Analyses: The WAAVES Program

James M. Reno[a,e], **Bryan Marker**[b], **Lawrence K. Cormack**[c], **Timothy Schallert**[a,e], and **Christine L. Duvauchelle**[d,e,*]

James M. Reno: jreno@utexas.edu; Bryan Marker: bamarker@cs.utexas.edu; Lawrence K. Cormack: cormack@utexas.edu; Timothy Schallert: schallert@psy.utexas.edu

[a]The University of Texas at Austin, Department of Psychology, Behavioral Neuroscience, 108 E. Dean Keeton Avenue, Stop A8000, Austin, TX 78712-1043, USA

[b]The University of Texas at Austin, Department of Computer Science, 2317 Speedway, Stop D9500, Austin, TX 78712, USA

[c]The University of Texas at Austin, Department of Psychology, Cognitive Systems and Center for Perceptual Systems, 108 E. Dean Keeton Avenue, Stop A8000, Austin, TX 78712-1043, USA

[d]The University of Texas at Austin, College of Pharmacy, Division of Pharmacology and Toxicology, 2409 University Avenue, Stop A1915, Austin, TX, 78712, USA

[e]Waggoner Center for Alcohol & Addiction Research, The University of Texas at Austin, 2500 Speedway, Stop A4800, Austin, TX 78712, USA

Abstract

Background—Human emotion is a crucial component of drug abuse and addiction. Ultrasonic vocalizations (USVs) elicited by rodents are a highly translational animal model of emotion in drug abuse studies. A major roadblock to comprehensive use of USV data is the overwhelming burden to attain accurate USV assessment in a timely manner. One of the most accurate methods of analyzing USVs, human auditory detection with simultaneous spectrogram inspection, requires USV sound files to be played back 4% normal speed.

New Method—WAAVES (WAV-file Automated Analysis of Vocalizations Environment Specific) is an automated USV assessment program utilizing MATLAB's Signal and Image Processing Toolboxes in conjunction with a series of customized filters to separate USV calls from background noise, and accurately tabulate and categorize USVs as flat or frequency-modulated (FM) calls. In the current report, WAAVES functionality is demonstrated by USV analyses of cocaine self-administration data collected over 10 daily sessions.

Results—WAAVES counts are significantly correlated with human auditory counts (r(48)=0.9925; p<0.001). Statistical analyses used WAAVES output to examine individual differences in USV responses to cocaine, cocaine-associated cues and relationships between USVs, cocaine intake and locomotor activity.

[*]Correspondence: Christine L. Duvauchelle, PhD, Associate Professor, The University of Texas at Austin, College of Pharmacy, Division of Pharmacology and Toxicology, 2409 University Avenue, Stop A1915, Austin, TX 78712, Tel: 512-471-1090, FAX: 512-475-6088, duvauchelle@mail.utexas.edu.

Comparison with Existing Method—WAAVES output is highly accurate and provides tabulated data in approximately 0.4% of the time required when using human auditory detection methods.

Conclusions—The development of a customized USV analysis program, such as WAAVES streamlines USV assessment and enhances the ability to utilize USVs as a tool to advance drug abuse research and ultimately develop effective treatments.

Keywords

Addiction; Craving; Cocaine; Rat; Self-administration; USV

1. Introduction

Emotion is a crucial component of drug use, addiction and withdrawal. Exhilarating effects of cocaine invites initial drug use, cues associated with cocaine trigger emotional yearning (e.g., "cocaine craving") for the drug, even after long periods of drug abstinence (Childress et al., 1999), and pervasive negative affect aggravates dependent individuals during drug withdrawal (Koob and Le Moal, 1997; Kreek and Koob, 1998). Preclinical research to understand drug abuse disorders and devise optimal treatment strategies could be greatly facilitated by monitoring progressive changes in emotionality related to drug use and drug-associated cues. It is well established that rodents emit ultrasonic vocalizations (USVs) to communicate emotional status (Brudzynski, 2009; Burgdorf et al., 2008; Knutson et al., 1999). Recently, the utility of USVs as indices of motivational states associated with natural (Brudzynski, 2009; Burgdorf et al., 2008) and drug reinforcement (Barker et al., 2010; Ma et al., 2010; Maier et al., 2012; Maier et al., 2010) has expanded beyond simple affirmation of positive and negative responses to stimuli. For instance, studies from our laboratory and others have demonstrated that USVs can serve as indices of cocaine-associated learning in a drug-free environment (Ma et al., 2010; Maier et al., 2010; Meyer et al., 2012) and reveal emotional reactions that are enhanced in response to cocaine and cocaine-associated environments after brief abstinence, and decreased after chronic cocaine self-administration experience (Maier et al., 2012; Maier et al., 2010). Though the ability to target crucial stages during the development of drug abuse disorders is attainable by tracking cocaine-associated USVs, continuous assessment of USVs through long-term drug abuse experiments is not without drawbacks.

USV assessment has become increasingly practical with the development of sensitive, high resolution microphones that can detect and discriminate variations in sound within ultrasonic frequencies (Specht, 2004). Precise detection and quantification of characteristics unique to USVs, such as bandwidth, frequency, amplitude, duration and inter-call interval, allows for thorough analyses of individual calls and categorization of discrete call types (Wright et al., 2010). Since USV occurrence can be observed in real-time, the functional significance of unique USV call types may be revealed by further studies linking certain USVs with ongoing and impending behaviors.

USVs can provide added insight to the study of drug abuse disorders. By monitoring USVs before, during and after drug intake, cocaine associative learning and progressive changes in affect accompanying transition to excessive drug use can be examined (Ahrens et al., 2009; Ma et al., 2010; Maier et al., 2012; Maier et al., 2010). In addition, individual differences in affective responses to drugs and drug-paired environments can be discerned by comparing USVs across subjects (Ahrens et al., 2009). Yet, a major roadblock limiting USV data utilization is the overwhelming burden to attain accurate USV assessment in a timely manner. For instance, a highly reliable method of USV quantification is obtained through human visual inspection of spectrogram files in conjunction with auditory confirmation.

This method assures that visualized sound objects can be further distinguished as USVs or noise. For humans to be able to detect sounds produced within the range of ultrasonic frequencies of interest (e.g., usually 30–100 kHz), playback of USV sound files must be slowed to 4% of its actual speed. In other words, since one second of data would take 20 seconds to play back at a distinguishable speed, a 10-min file would require at least 4 hours to tabulate. Considering that drug self-administration sessions are typically at least 1-hr in duration, unanalyzed data sets tend to accumulate rapidly. The inevitable lag between data collection and analyses may be part of the rationale to report small subsets of USV data rather than complete drug administration sessions (e.g., first 30 mins of 2 hrs; 3 out of 20 mins) (Mahler et al., 2013; Wright et al., 2012).

There are available software programs for automated USV assessment (e.g. Avisoft's SASLAB Pro; Avisoft Bioacoustics, Berlin, Germany) that can tabulate sounds occurring within certain ultrasonic frequency ranges. However, programs designed for naturalistic settings are not specifically suited for experimental environments common to pre-clinical drug abuse studies, such as operant chambers equipped with levers, drug swivels and counter-balances. As a result, background sounds, such as those produced by animal activity, contact between any objects within the test chamber, or any noise-producing event within the ultrasonic microphone detection range may be automatically tabulated as USV calls. Thus, in creating a program to accurately assess USVs, it is imperative to identify distinguishing characteristics to separate USV calls from noise. In the process of developing an automated program, we found that background noise varies greatly between test environments and experimental conditions. Therefore, the parameters that we report here are unique to our operant chamber environment (see Materials and Methods below).

The automated program we developed, "WAAVES" (WAV-file Automated Analysis of Vocalizations Environment Specific) utilizes MATLAB's (The Mathworks, Inc., Natick, MA) *Signal* and *Image Processing Toolboxes* to import USV files and identify signals meeting base USV call criterion. The potential USV calls are then filtered through a series of requirements that separates USV calls from noise.

To demonstrate the functionality of WAAVES, USVs in the 50–55 kHz (flat and frequency-modulated, or FM types) and 22–28 kHz ranges were assessed in a group of rats self-administering cocaine in a self-administration and *in vivo* microdialysis experiment. Using the WAAVES findings, animals were categorized into High and Low USV Caller groups. Statistical analyses were then used to examine relationships between individual differences in 50–55 kHz USV responses to cocaine and cocaine-associated cues, with self-administered cocaine intake and locomotor activity data. These data have not been previously published.

2. Materials and Methods

2.1.1 Animals and Surgery

Male Sprague-Dawley rats were received at 5 weeks of age and handled for at least 2 weeks prior to the start of the experiment. Animals were trained to press a lever with sucrose pellets as reinforcement. All animals underwent jugular catheterization and stereotaxic surgery as previously described (Depoortere et al., 1993; Feduccia et al., 2010; Maier et al., 2008).

2.1.2 Apparatus, Groups and Experimental Procedure

The apparatus consisted of a single retractable-lever operant chamber (28x22x21 cm) located within a sound-attenuating box (Med Associates, St. Albans, VT). The lever was located on the right wall in the back corner with a stimulus light above it. An ultrasonic microphone (UltraSoundGate CM16/CMPA, Avisoft Bioacoustics, Berlin, Germany) was

located in the center panel of the right wall, 1 cm from the chamber ceiling and a maximum distance of 36 cm from any point in the chamber.

Within these chambers, all animals (n=15) participated in 10-day (5 days/week) cocaine self-administration sessions. Daily sessions began with a 10-min pre-drug ("cocaine anticipatory") interval within the darkened operant chamber. During this time, the lever was retracted and cocaine was not available. Immediately following the pre-drug interval, chamber lights were illuminated and a one-hour cocaine self-administration session commenced. The animals were presented a lever (accompanied by the illumination of a cue light) for 30 seconds three times per ten minutes at variable intervals (18 presentations total). The first presentation was at the beginning of each 10-minute period and the second two were at 120, 180 or 240 seconds after the previous presentation (interval was randomly selected without replacement during each 10-minute period). A lever response during cued cocaine availability intervals resulted in a single cocaine injection (0.75 mg/kg in 0.1 ml; 6-second infusion duration) accompanied by continuous flashing of the cue light. The lever was withdrawn between cocaine access periods.

2.2 WAAVES' USV Quantification

2.2.1 Identifying Objects in Sound Files Using MATLAB—Using MATLAB's (The Mathworks, Inc., Natick, MA) *Signal Processing* and *Image Processing Toolboxes,* WAAVES reads sound files, creates a spectrogram, converts the spectrogram to grayscale and identifies objects within the image for closer inspection. It should be noted that the current program is written to analyze WAV-files, however any format of sound file can be analyzed with minor adjustments.

2.2.2 Developing Separation Parameters—At each phase of developing the WAAVES program, WAAVES output was compared to original sound files and manual adjustments were made to the program as needed. Call identification successes and failures (failures defined as missing calls, incorrect call counts and/or not correctly identifying the beginning and ending of a call) revealed patterns that assisted in fine-tuning the WAAVES program. Repeating this procedure and manipulating the exact value of each criterion in concert with additional criteria allowed for precise separation of calls from noise. Separate call criteria were developed for 50–55 kHz versus 22–28 kHz calls (discussed in detail below). It is important to note that each distinct test environment (e.g. operant chamber, open field, home cage, etc.) requires customized separation criteria. The guidelines established for WAAVES versions reported here were specifically customized to our cocaine self-administration apparatus and methodology. This version of the WAAVES program extends to all of our operant chambers (e.g. new criteria did not need to be developed for each individual chamber). Therefore it is likely that a similarly customized program for USV analyses would be equally reliable across identical test environments of other varieties (e.g. home cage, open field, etc.).

2.2.3 50–55 kHz Calls Criteria—Calls detected within the frequency range of 30–90 kHz were considered 50–55 kHz (e.g., "positive affect") USVs. Background noise can occur at all frequencies, but can be differentiated from calls by identifying elements unique to noise. The process starts by identifying "objects" (e.g., potential USV calls) within sound files using MATLAB's (The Mathworks, Inc., Natick, MA) *Signal and Image Processing Toolboxes*. The first elements defined as noise are: 1) elements that span 5 or more consecutive frequencies within a 0.5 ms period, and 2) elements co-occurring with another sound element between 20–30 kHz (e.g., usually a 22–28 kHz call). This second criterion was added to the 50–55 kHz WAAVES program based our observations that, within the closed operant chamber, high decibel 22–28 kHz USVs can reverberate at a higher

frequency (e.g., between 40–60 kHz) and thus appear as separate sound elements within the 50–55 kHz range. 3) Elements exceeding a range of 60 kHz per 0.5 ms (max frequency – min frequency) and/or with a mean frequency greater than 90 kHz were filtered out as noise. Since we observed that background noise consists largely of extremely brief elements, also filtered out as noise were 4) sound elements less than 2.5 ms in duration.

After the initial screening of noise, all surviving elements were sorted into individual calls depending on their temporal location to other elements. For instance, to be considered separate calls, a 10-ms "hold time," (e.g., the period of time between individual calls) was required. If the inter-element space was less than 10 ms, the elements were considered units of a single call. Another characteristic of USVs is call duration. Call duration can be set according to experimenter preference, but for this report, minimum call durations were set at 5 ms and 20 ms.

Comparisons between spectrograms and automated USV assessment output need to be repeatedly performed throughout development of an automated program to assure that noise elements are not included in USV call assessment. For example, in our operant chambers, loud background noise is often detected as a large sound element spanning a wide range of frequencies, including those in which 50–55 kHz USVs would occur. By viewing the spectrogram and comparing the automated output to the observed calls, specific distinctions between calls and characteristics of noise generated within our experimental apparatus were determined and adjustments to programming parameters were executed. In this case, to avoid counting these noise elements as calls, the final separation criterion considered the change in mean frequency in 0.5 ms steps across the entire call and required the median value to be greater than zero and less than two in order to be tabulated as a USV call. This final criterion is effective as background noise either fluctuates greatly in mean frequency, or not at all.

At this point in WAAVES, positive "50–55 kHz" USVs are separated from background noise and characteristics of interest can be derived (see Figure 1). For instance, in this WAAVES version frequency-modulated (FM) USV calls were defined as those varying more than 5 kHz in their maximum frequency over the entire call (e.g., maximum frequency at each 0.5 ms time point was collected and if the maximum minus the minimum of these values was 5 or above the call was defined as an frequency-modulated (FM) call). Flat calls were defined as those not meeting the FM criterion.

2.2.4 22–28 kHz Calls Criteria—Low frequency 22–28 kHz calls were tabulated when occurring within the frequency range of 20–30 kHz. Maximum frequency range of detected elements was 10 kHz and the mean frequency was required to be less than 30 kHz. Similar to the positive WAAVES program, objects less than 2.5 ms were considered to be noise. Minimum call duration was 200 ms with a hold time of 100 ms between calls. Low frequency calls require fewer separation criteria to distinguish calls from background noise, due to their typically consistent features, including long duration, uniform frequency and decibel level.

3. Results

3.1 Statistical Analyses

Pearson's correlation tests were used to determine the degree of similarity between Human Auditory/Visual counts and WAAVES USV counts, as well as to examine various relationships between behavioral indices (e.g., 50–55 kHz FM USVs, cocaine intake and locomotor activity) during Cocaine Anticipatory (e.g., the 10-min drug-free intervals prior to

cocaine access) and Cocaine Self-Administration Sessions (e.g., the 60-min cocaine access period) in both High and Low USV Caller groups.

Two-way repeated measures ANOVAs (Group X Days) compared 22–28 and 50–55 kHz USVs and locomotor activity in High and Low USV Caller groups during Cocaine Anticipatory (drug-free) intervals, and USVs, locomotor activity and cocaine intake during Cocaine Self-Administration Sessions. Posthoc comparisons (Fisher's Least Significant Difference (LSD) tests) were performed when significant interactions were detected.

3.2 Human Auditory/Visual and WAAVES USV Output Results

Fifty USV sound files (1-min duration) recorded during cocaine self-administration sessions were analyzed using the WAAVES program and by research assistants trained to recognize flat and FM 50 kHz USVs during slowed playback and simultaneous spectrogram inspection. Pearson's analyses of the counts from each method showed significant linear relationships between WAAVES output and human auditory/visual detection (5ms and up: $r(48)=0.9925$ and 20ms and up: $r(48)= 0.9897$; $p<0.001$ for both, see Figure 2). WAAVES analyses were used for all USV assessments reported below.

3.3 High and Low USV Caller Group Assignments

To examine the implications of individual differences in USV responses to cocaine and cocaine-associated cues, the number of 50–55 kHz USV calls (both flat and FM) emitted by each rat (n=15) across all ten cocaine self-administration sessions were sorted to determine the range of total USV calls for all rats. The number of 50–55 kHz USVs emitted by one rat was three standard deviations above the overall group mean, so these data were eliminated from further analyses. Data from the remaining rats (n=14) were then equally divided into "High USV Caller" (HC; n=7) and "Low USV Caller" (LC; n=7) groups.

3.4 Two-way ANOVA Results

50–55 kHz FM calls are considered to be specifically associated with drug-associated positive affect (Ahrens et al., 2009; Knutson et al., 1999). Therefore, two-way ANOVAs were performed on 50–55 kHz FM USVs. Statistical significance between High and Low USV Callers and 50–55 kHz USV durations of either 5 or 20 ms and up were virtually identical. To avoid repetitive data reporting, statistical results were calculated using WAAVES outputs of 50–55 kHz FM USV durations of 20 ms and up.

Cocaine Intake - Cocaine Self-Administration Sessions—The number of cocaine injections during Cocaine Self-Administration Sessions did not significantly differ between the groups designated as High and Low USV Callers ($F(1,12)=1.23$; n.s.). In addition, there were no Group by Day Interaction effects ($F(9,108)=1.527$; n.s.), but significant Day effects ($F(9,108)=10.11$; $p<0.0001$) were observed (see Figure 3).

Locomotor Activity – Cocaine Anticipatory Intervals and Cocaine Self-Administration Sessions—Locomotor activity did not differ between High and Low USV Caller Groups during Cocaine Anticipatory intervals ($F(1,12)=2.976$; n.s.) or during Cocaine Self-Administration Sessions ($F(1,12)=0.0805$; n.s.; see Figure 4). Significant Day effects during Cocaine Self-Administration Sessions were observed ($F(9,108)=10.17$; $p<0.0001$), but no other significant differences were detected.

50–55 kHz FM USVs: Cocaine Anticipatory Intervals—Significant differences between High and Low USV Callers were revealed in Group ($F(1,12)=4.518$; $p=0.05$), Day ($F(9,108)=3.259$; $p=0.0015$) and Interaction effects ($F(9,108)=2.335$; $p=0.0191$). Post hoc (LSD tests) showed High Callers produced significantly more FM calls than Low Callers

specifically during the last ½ of the experiment (e.g., on Days 6, 8, 9, and 10; p<0.01, data not shown)

50–55 kHz FM USVs – Cocaine Self-Administration Sessions—During Cocaine Self-Administration Sessions, significant Day ($F_{(9,108)}=2.613$; p=0.009) and Interaction effects ($F_{(9,108)}=2.234$; p=0.025) between High and Low USV callers were observed. Overall Group differences approached significance ($F_{(1,12)}=4.159$; p =0.064, see Figure 5).

22–28 kHz USVs – Cocaine Anticipatory Intervals and Cocaine Self-Administration Sessions—A small number of 22–28 kHz USVs (USV duration set at 200 ms and above) were tabulated during experimental sessions. Two-way ANOVAs performed on 22–28 kHz USVs demonstrated no differences during Cocaine Anticipatory intervals (Group: $F_{(1,12)}=0.7955$, Day: $F_{(9, 108)}=0.9873$, Interaction: $F_{(9, 108)}=0.9976$; all n.s.) or during Cocaine Self-Administration Sessions (Group: $F_{(1,12)}=0.8923$, Day: $F_{(9,108)}=1.002$, Interaction $F_{(9, 108)}=0.9582$; all n.s., data not shown).

3.5 Pearson's Correlation Results

Pearson' correlation were performed to determine linear relationships between 50–55 kHz FM USVs, self-administered cocaine intake and locomotor activity in High and Low USV Caller groups. Significant relationships are listed below:

Cocaine Anticipatory FM USVs and Cocaine Self-Administration—For High USV callers, a significant positive linear relationship was detected between the number of FM USVs elicited during the Cocaine anticipatory interval and the number of self-administered cocaine infusions during Cocaine Self-Administration Sessions ($r_{(68)}=0.3450$; p<0.01). For Low USV Callers, the number of cocaine anticipatory FM USVs and cocaine self-administration levels were not significantly correlated ($r_{(68)}=0.1227$; n.s., see Figure 6A).

Cocaine Anticipatory FM USVs and Cocaine Self-Administration USVs—There was a significant linear relationship between FM USVs elicited during Cocaine Anticipatory intervals and FM USVs elicited during Cocaine Self-Administration in both High and Low USV Caller groups (HC: $r_{(68)}=0.8518$; p<0.001, LC: $r_{(68)}=0.2354$; p<0.05, see Figure 6B).

Cocaine Self-Administration and Cocaine Self-Administration USVs—In High and Low USV Callers, linear relationships between the number of self-administered cocaine infusions and FM USVs elicited during Cocaine Self-Administration Sessions were significant (HC: $r_{(68)}= 0.2920$; p<0.01, LC: $r_{(68)}=0.2663$; p<0.05, data not shown).

Cocaine Anticipatory FM USVs and Cocaine Anticipatory Locomotor Activity—During the Cocaine Anticipatory interval, locomotor activity counts and FM USVs were significantly correlated for High and Low USV Callers (HC: $r_{(68)}=0.3448$; p<0.01, LC: $r_{(68)}=0.3261$; p<0.01, see Figure 6C).

4. Discussion

Ultrasonic vocalization data can provide pre-clinical high-resolution "real-time" surveys of affective changes over the course of drug experience. With the development of the WAAVES program, comprehensive USV data analyses can be rapidly achieved. For example, to assess USVs using human auditory detection and simultaneous spectrogram inspection, the playback time of the WAV files running at 4% original speed would be 29.16 hours (1750 minutes) for a single experimental session (e.g., 10-min drug free interval + 60-

min of cocaine availability). In contrast, to tabulate the same amount of data using WAAVES takes an average of 5.53 minutes. For experiments that record substantial amounts of USV data, the considerable time and effort needed to develop a customized method for automated assessment is a wise investment. The current report further demonstrates the utility of USVs and the WAAVES program by ascertaining individual differences in 50–55 kHz FM USV responses to cocaine and determining linear relationships between FM USVs, cocaine intake and locomotor activity.

High versus Low USV Callers differed significantly in the number of 50–55 kHz FM USVs elicited during Cocaine Anticipatory intervals and during Cocaine Self-Administration Sessions group differences approached significance. Yet, reinforced cocaine response rates, cocaine-stimulated and cocaine-conditioned hyperlocomotor activity; measures commonly considered to be standard indices of cocaine motivation (Horger et al., 1990; Panlilio and Schindler, 1997; Piazza et al., 1989), did not vary significantly between High and Low USV Caller groups. These findings demonstrate that 50–55 kHz FM USVs and traditional behavioral measures of drug reward and/or reinforcement are not similarly influenced by cocaine experience.

Since 50–55 kHz USVs lack equivalence with behavioral assessments customarily associated with cocaine self-administration, some may consider USVs inappropriate for drug abuse research. However, this view can be easily discounted when determining linear relationships between cocaine-conditioned 50–55 kHz FM USVs and cocaine-stimulated behaviors. For example, correlational analyses revealed that the number of USVs elicited in a cocaine-associated environment (e.g., Cocaine Anticipatory USVs) in High USV Callers is significantly correlated with cocaine intake once cocaine becomes available. In other words, future high cocaine intake may be predicted by robust USV responses elicited within a cocaine-conditioned environment. If 50–55 kHz FM USVs signal positive affective reactions, as is consistent with prevailing notions on the matter (Burgdorf et al., 2008), the translational relevance of these findings are evident; corresponding with clinical reports of positive anticipation of cocaine influencing intake behavior (Browning et al., 2011; Weiss et al., 2003) and cocaine relapse in response to cue-induced cocaine craving (O'Brien et al., 1998). In addition, the significant linear relationship between 50–55 kHz anticipatory FM USVs and drug-induced FM USVs observed in both High and Low USV Callers, suggest that the level of FM USVs elicited in anticipation of cocaine may predict the magnitude of positive affect induced by self-administered cocaine. This extends previous findings that initial drug-induced USVs are predictive of future drug consumption (Browning et al., 2011), demonstrating now that USVs made in a drug-free state (e.g. 50–55 kHz anticipatory FM USVs) can be predictive of subsequent drug-taking and drug-induced positive affect.

To move the field of USV analyses forward, categorization of USV call features needs to be standardized. Replacing subjective judgments of human analyzers with uniform measurements can be achieved with an automated USV analysis program and can be applied to various USV characteristics of interest. For example, USV duration is a trait associated with different USV categories (Brudzynski, 2009). However, when USVs are assessed through human visual and auditory observation, it is difficult to accurately or consistently tally USV duration in all calls. Indeed, though published literature describes 50–55 kHz USVs durations as above 20 ms (Brudzynski, 2009), WAAVES output showed a marginally stronger (not statistically different) correlational coefficient with human assessment when USV duration parameters were set at 5 ms compared to 20 ms (e.g., $r(48) = 0.9925$ versus 0.9897). As we learn more about USVs and components of calls that emerge as behaviorally and/or pharmacologically relevant, automated analyses programs can be updated and previously recorded sound files can be readily adapted for changing parameters and new characteristics of interest.

The development of a customized USV analysis program (WAAVES) has streamlined the ability to guide ongoing experiments with up-to-the-minute USV data, evaluate USV data across long-term drug administration experiments and categorize USV call types. These factors alone make an automated USV analysis program, such as WAAVES, an important tool for drug abuse research. Although an automated USV analysis program needs to be customized for different testing environments, in the end, it is possible to achieve automated outputs of USV counts nearly identical to human assessment measures in a fraction of the time.

As a major component of drug abuse and addiction, alterations in emotion accurately modeled in animals is a major step forward for drug abuse research. Comprehensive USV studies have the potential to provide insight into changes in emotional responses that accompany and/or induce progression from initial to compulsive drug use, abuse and relapse. Automated programs like WAAVES address the bottle-neck in data analyses by practically eliminating the time-intensive nature of USV data analyses. The linear relationships between USVs and observable cocaine-motivated behaviors reported here confirm the utility of USVs as a highly translational and unique measure of cocaine-induced reward and cocaine-conditioning, and a potential predictive measure of future drug-taking behavior.

Acknowledgments

References

Ahrens AM, Ma ST, Maier EY, Duvauchelle CL, Schallert T. Repeated intravenous amphetamine exposure: rapid and persistent sensitization of 50-kHz ultrasonic trill calls in rats. Behav Brain Res. 2009; 197:205–9. [PubMed: 18809437]

Barker DJ, Root DH, Ma S, Jha S, Megehee L, Pawlak AP, West MO. Dose-dependent differences in short ultrasonic vocalizations emitted by rats during cocaine self-administration. Psychopharmacology (Berl). 2010; 211:435–42. [PubMed: 20571780]

Browning JR, Browning DA, Maxwell AO, Dong Y, Jansen HT, Panksepp J, Sorg BA. Positive affective vocalizations during cocaine and sucrose self-administration: a model for spontaneous drug desire in rats. Neuropharmacology. 2011; 61:268–75. [PubMed: 21530553]

Brudzynski SM. Communication of adult rats by ultrasonic vocalization: biological, sociobiological, and neuroscience approaches. ILAR J. 2009; 50:43–50. [PubMed: 19106451]

Burgdorf J, Kroes RA, Moskal JR, Pfaus JG, Brudzynski SM, Panksepp J. Ultrasonic vocalizations of rats (Rattus norvegicus) during mating, play, and aggression: Behavioral concomitants, relationship to reward, and self-administration of playback. J Comp Psychol. 2008; 122:357–67. [PubMed: 19014259]

Childress AR, Mozley PD, McElgin W, Fitzgerald J, Reivich M, O'Brien CP. Limbic activation during cue-induced cocaine craving. Am J Psychiatry. 1999; 156:11–8. [PubMed: 9892292]

Depoortere RY, Li DH, Lane JD, Emmett-Oglesby MW. Parameters of self-administration of cocaine in rats under a progressive-ratio schedule. Pharmacol Biochem Behav. 1993; 45:539–48. [PubMed: 8332614]

Feduccia AA, Kongovi N, Duvauchelle CL. Heat increases MDMA-enhanced NAcc 5-HT and body temperature, but not MDMA self-administration. Eur Neuropsychopharmacol. 2010; 20:884–94. [PubMed: 20888192]

Horger BA, Shelton K, Schenk S. Preexposure sensitizes rats to the rewarding effects of cocaine. Pharmacol Biochem Behav. 1990; 37:707–11. [PubMed: 2093176]

Knutson B, Burgdorf J, Panksepp J. High-frequency ultrasonic vocalizations index conditioned pharmacological reward in rats. Physiol Behav. 1999; 66:639–43. [PubMed: 10386908]

Koob GF, Le Moal M. Drug abuse: hedonic homeostatic dysregulation. Science. 1997; 278:52–8. [PubMed: 9311926]

Kreek MJ, Koob GF. Drug dependence: stress and dysregulation of brain reward pathways. Drug Alcohol Depend. 1998; 51:23–47. [PubMed: 9716928]

Ma ST, Maier EY, Ahrens AM, Schallert T, Duvauchelle CL. Repeated intravenous cocaine experience: development and escalation of pre-drug anticipatory 50-kHz ultrasonic vocalizations in rats. Behav Brain Res. 2010; 212:109–14. [PubMed: 20382187]

Mahler SV, Moorman DE, Feltenstein MW, Cox BM, Ogburn KB, Bachar M, McGonigal JT, Ghee SM, See RE. A rodent "self-report" measure of methamphetamine craving? Rat ultrasonic vocalizations during methamphetamine self-administration, extinction, and reinstatement. Behav Brain Res. 2013; 236:78–89. [PubMed: 22940018]

Maier EY, Abdalla M, Ahrens AM, Schallert T, Duvauchelle CL. The missing variable: ultrasonic vocalizations reveal hidden sensitization and tolerance-like effects during long-term cocaine administration. Psychopharmacology (Berl). 2012; 219:1141–52. [PubMed: 21870038]

Maier EY, Ahrens AM, Ma ST, Schallert T, Duvauchelle CL. Cocaine deprivation effect: cue abstinence over weekends boosts anticipatory 50-kHz ultrasonic vocalizations in rats. Behav Brain Res. 2010; 214:75–9. [PubMed: 20470830]

Maier EY, Ledesma RT, Seiwell AP, Duvauchelle CL. Diazepam alters cocaine self-administration, but not cocaine-stimulated locomotion or nucleus accumbens dopamine. Pharmacol Biochem Behav. 2008; 91:202–7. [PubMed: 18691612]

Meyer PJ, Ma ST, Robinson TE. A cocaine cue is more preferred and evokes more frequency-modulated 50-kHz ultrasonic vocalizations in rats prone to attribute incentive salience to a food cue. Psychopharmacology (Berl). 2012; 219:999–1009. [PubMed: 21833503]

O'Brien CP, Childress AR, Ehrman R, Robbins SJ. Conditioning factors in drug abuse: can they explain compulsion? J Psychopharmacol. 1998; 12:15–22. [PubMed: 9584964]

Panlilio LV, Schindler CW. Conditioned locomotor-activating and reinforcing effects of discrete stimuli paired with intraperitoneal cocaine. Behav Pharmacol. 1997; 8:691–8. [PubMed: 9832954]

Piazza PV, Deminiere JM, Le Moal M, Simon H. Factors that predict individual vulnerability to amphetamine self-administration. Science. 1989; 245:1511–3. [PubMed: 2781295]

Specht R. Compact and user-friendly ultrasound acquisition systems optimized for field recording. An Acad Bras Cienc. 2004; 76:453–4. [PubMed: 15258665]

Weiss RD, Griffin ML, Mazurick C, Berkman B, Gastfriend DR, Frank A, Barber JP, Blaine J, Salloum I, Moras K. The relationship between cocaine craving, psychosocial treatment, and subsequent cocaine use. Am J Psychiatry. 2003; 160:1320–5. [PubMed: 12832248]

Wright JM, Dobosiewicz MR, Clarke PB. The role of dopaminergic transmission through D1-like and D2-like receptors in amphetamine-induced rat ultrasonic vocalizations. Psychopharmacology (Berl). 2012

Wright JM, Gourdon JC, Clarke PB. Identification of multiple call categories within the rich repertoire of adult rat 50-kHz ultrasonic vocalizations: effects of amphetamine and social context. Psychopharmacology (Berl). 2010; 211:1–13. [PubMed: 20443111]

Highlights

- Automated program to analyze ultrasonic vocalization (USV) sound files was developed

- Highly significant correlation between program output and human assessments

- Program completes analyses in approximately 0.4% of time compared to manual methods

- USVs shown as useful index for detecting cocaine-associated individual differences

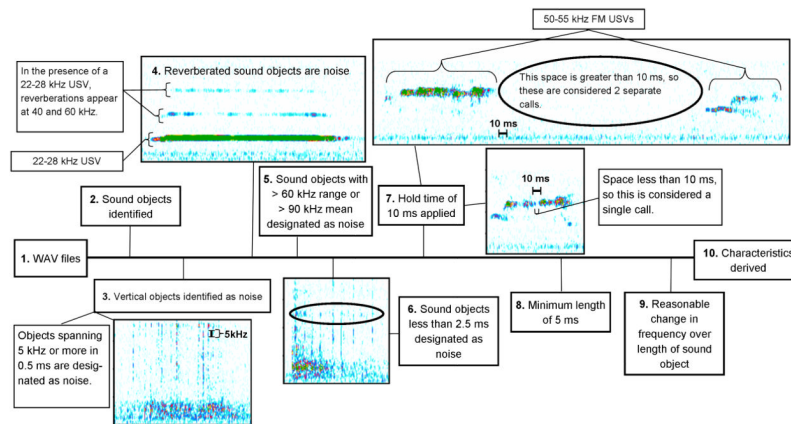- Statistical analyses reveal cocaine-associated USVs predictive of cocaine intake

Figure 1. WAAVES Program Flowchart
The order of separation criteria is an important factor when developing an automated analysis program such as WAAVES. **1.** Sound files (.wav) are read into the program. **2.** Sound objects are identified within the files for closer inspection. **3.** Sound objects spanning 5 kHz and above consecutive frequencies in 0.5 ms are considered noise **4.** Within the experimental apparatus, noise generated by reverberations from high decibel 22–28 kHz USVs often appears at 40 and 60 kHz **5.** Sound objects defined as noise if frequency range > 60 kHz and mean frequency > 90 kHz **6.** Sound objects defined as noise if < 2.5 ms **7.** Inter-call interval at least 10 ms for call count > 1 **8.** Call duration setting (e.g., 5 or 20 ms) **9.** Median value of change in mean frequency in 0.5 ms steps was required to be > 0 and < 2 kHz. **10.** At this point calls have been successfully separated from noise and further USV subcategories can be created (e.g. flat versus frequency-modulated (FM)).
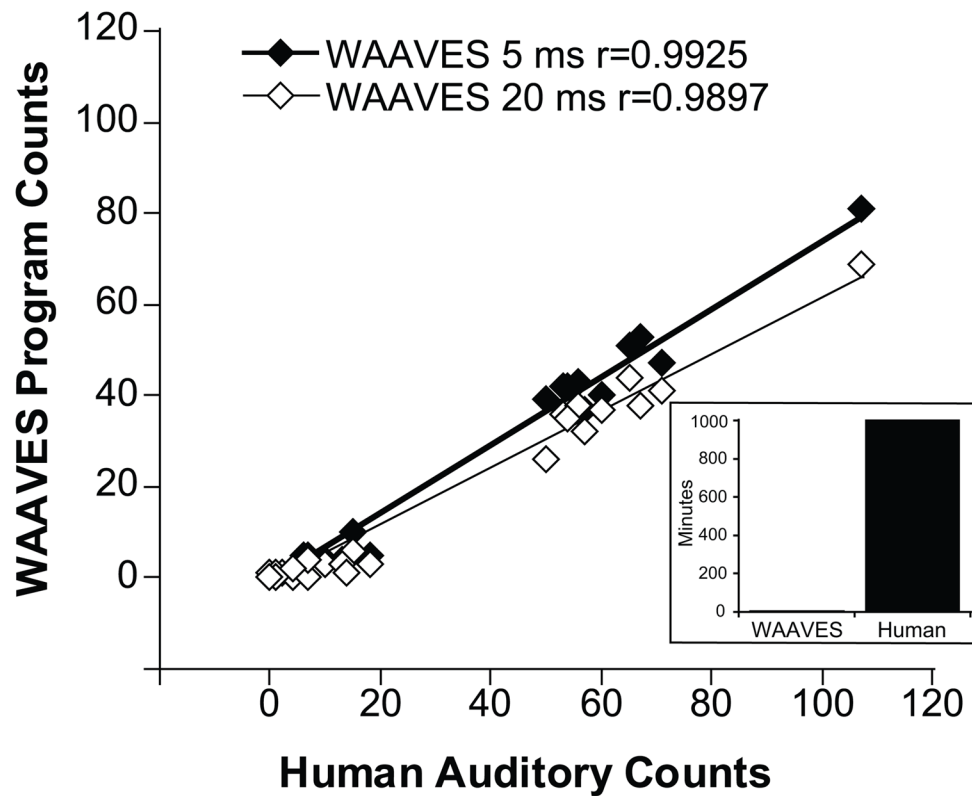
Figure 2. Comparison between WAAVES output and human auditory counts
The WAAVES outputs using either 5 or 20 ms USV duration criteria were significantly correlated with the number of USVs tabulated using human auditory confirmation (p<0.001 for both). Inset bar graph displays the average time (mins) to analyze 50 1-min USV data files data using WAAVES (≈3.75 mins) compared to the amount of time needed for playback (≈1000 mins) of the same sound files at a speed allowing for human auditory detection.
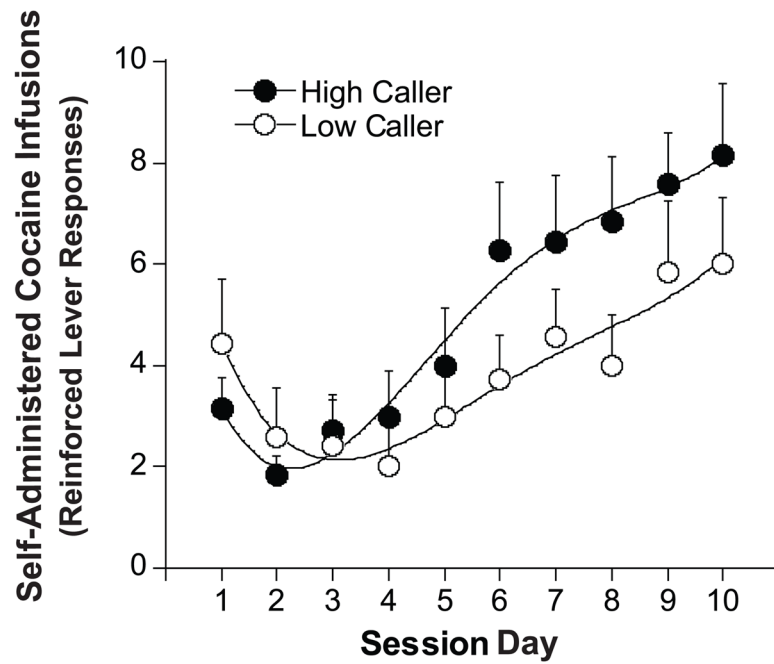
Figure 3. Cocaine Intake in High and Low USV Callers
The number of self-administered cocaine infusions (0.75 mg/kg/inf) did not significantly differ between High and Low USV Callers. Data are daily mean (+SEM) cocaine infusions shown with a fourth degree polynomial line fit.
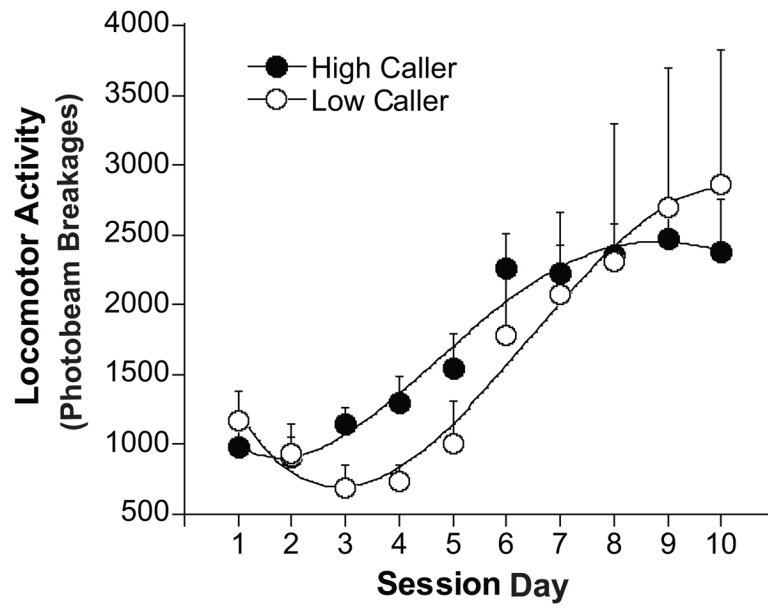
Figure 4. Cocaine-stimulated locomotor activity in High and Low USV Callers
Cocaine-stimulated locomotor activity was similar in High and Low USV Caller groups.
Data are daily mean (+SEM) locomotor activity levels (number of photobeam breakages)
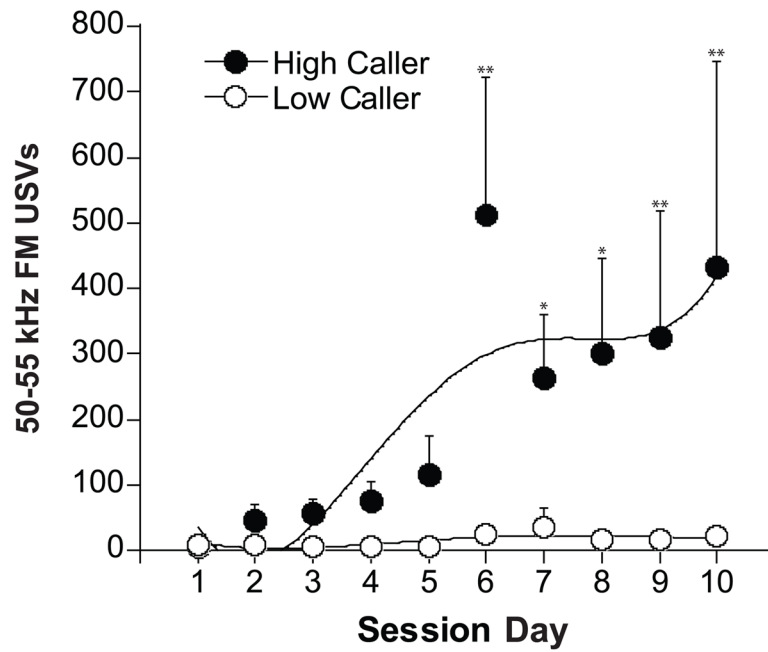shown with a fourth degree polynomial line fit.

Figure 5. 50–55 kHz FM USVs in High and Low USV Callers
During access to self-administered cocaine, 50–55 kHz FM USVs were significantly greater in the High, compared to the Low USV Caller group. Data are daily mean (+SEM) 50–55 kHz USVs shown with a fourth degree polynomial fit. *, ** = significant difference between daily group means at p<0.05 and <0.01, respectively.
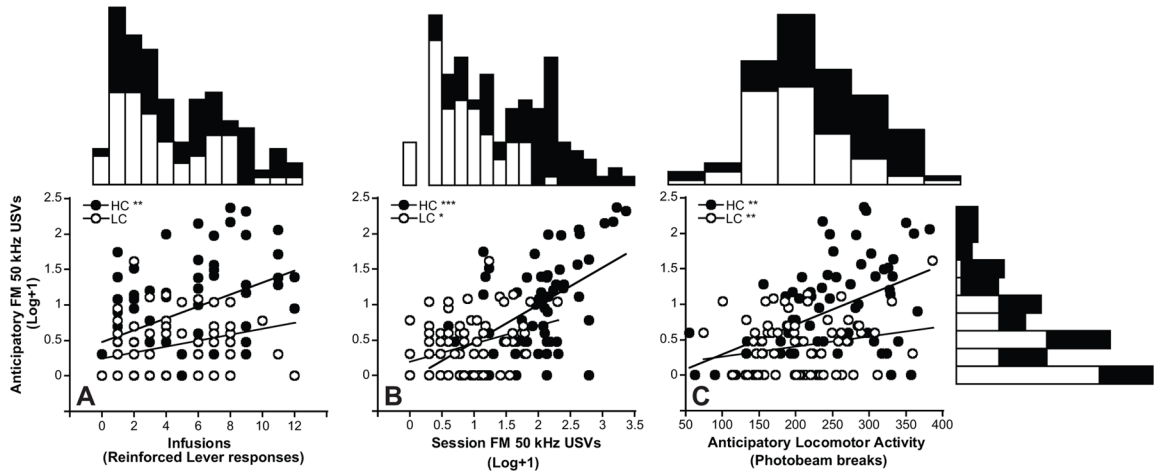
Figure 6. Significant Linear Relationships between Cocaine Anticipatory 50–55 kHz FM USVs and Self-Administered Cocaine, Cocaine-Induced 50–55 kHz FM USV, and Cocaine Anticipatory Locomotor Activity

50–55 kHz FM USV data was log-transformed (log+1) for graphic depiction only. Pearson's correlation tests were performed on raw data. Significant linear relationships were observed between Cocaine Anticipatory 50–55 kHz FM USVs and A) the number of self-administered cocaine infusions in the High USV Caller Group ($p<0.01$), B) the number of 50–55 kHz USVs elicited during Cocaine Self-Administration Sessions in both High and Low USV Caller groups ($p <0.001$ and 0.05, respectively), and C) the number of locomotor activity counts during Cocaine Anticipatory intervals for both High and Low USV Callers ($p <0.01$ for both; note log-transformed fit line for LC not reflecting significant linear relationship from raw data). Marginal histograms depict distribution of High and Low USV Callers across all measures.