Beyond dichotomy: ultrasonic vocalizations of rats convey more than sad and happy

Reo Wada [1#], Shiomi Hakataya [2#], Ryosuke O. Tachibana [3#], Tomoyo Isoguchi Shiramatsu [4], Tetsufumi Ito [5], Kouta Kanno [6], Riseru Koshiishi [1,7], Jumpei Matsumoto [5], Yumi Saito [7], Genta Toya [7,8], Shota Okabe [9**], Kazuo Okanoya [7*]

1. Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan.
2. University of the Ryukyus, Okinawa, Japan.
3. Human Information and Interaction Research Institute, National Institute for Advanced Industrial Science and Technology, Tsukuba, 305-8566, Japan.
4. Graduate School of Information Science and Technology, The University of Tokyo, Tokyo, Japan.
5. Faculty of Medicine, University of Toyama, Toyama, 930-0194, Japan.
6. Laboratory of Neuroscience, Course of Psychology, Department of Humanities, Faculty of Law, Economics and Humanities, Kagoshima University, Kagoshima, Japan.
7. Advanced Comprehensive Research Organization, Teikyo University, Tokyo, Japan.
8. Research Center for Advanced Science and Technology, The University of Tokyo
9. Jichi Medical University, Shimotsuke, Japan.

[#] equally contributed
** co-corresponding author (Shota Okabe)
* corresponding author (Kazuo Okanoya)

Author Contributions: RW, SH, and ROT wrote manuscripts. ROT performed data processing, visualization, and preparation for supplementary information. RW, SH, RK, and GT originally started the USV data survey project. All authors joined in the survey. SO and KO supervised the project. All authors were involved in revising and editing the draft.

Competing Interest Statement: The authors declare no competing interest.

Classification: Social Sciences/Psychological and Cognitive Sciences

Keywords: Ultrasonic vocalization; Frequency; Duration; Meta-analysis

Abstract

Ultrasonic vocalizations (USVs) in rats have been established as an ideal model for studying emotional expression because they reflect internal emotional states. Since the discovery of USVs in rats, it has gradually been revealed that two types of adult rats' USVs with distinct acoustic features (high-short and low-long) are closely linked to positive and negative affect, respectively. Thus, for many years, most studies have grouped USVs into dichotomous categories: 50-kHz (positive) USVs and 22-kHz (negative) USVs, using these numbers as labels rather than measured values. However, the diversity of rats' USVs may have been overlooked due to this 50- / 22-kHz dichotomous framework, as several recent studies claim to have found new categories in USVs outside the framework. Here, through a systematic survey of descriptions of rats' USVs in the past literature, we confirmed a non-negligible number of reports on USVs that are outside the dichotomous framework. Our results support the abundance of diverse USVs beyond the dichotomy. This extends the traditional association between acoustic features and emotional states, highlighting the importance of research on diverse vocalizations that might reflect complex affective states beyond the simple distinction between positive and negative.

Main Text

Introduction

Understanding animal emotions is of great interest to researchers across various disciplines, ranging from neuroscience to animal welfare. Rats (*Rattus norvegicus*), in particular, provide an ideal model for investigating emotional expression through their ultrasonic vocalizations (USVs), as these vocalizations reflect internal emotional states. Since the initial discovery of rat USVs (1), it has become increasingly evident that adult rats produce two distinct types of USVs, each linked to specific emotional contexts. Relatively high frequency with short-duration syllables ("high-short" calls; 45–80 kHz, 5–200 ms) are typically emitted during positive situations, whereas low and long syllables ("low-long" calls; 18–28 kHz, 500–2000 ms) characterize negative situations (2) (***Fig. 1A***).

Over the past two to three decades, rat USVs have attracted considerable attention across multiple fields, including medical, pharmacological, and behavioral research, as evidenced by a substantial increase in related publications (***Fig. 1B upper***). Consequently, rat USVs are widely used as indicators of emotional and internal states. Traditionally, researchers have grouped these vocalizations into the two categories corresponding to high-short and low-long calls, commonly referred to as "50 kHz" and "22 kHz" calls, respectively. This dichotomous classification has become predominant, with most studies using these two frequency categories as convenient labels rather than precise acoustic measurements (***Fig. 1B lower***). Despite attempts to document acoustic diversity within rat USVs by identifying additional subtypes (3, 4), the dichotomous framework remains prevalent. These subtypes continue to be viewed as mere subdivisions under the primary "50 kHz" and "22 kHz" categories, and there remains limited experimental evidence regarding their specific behavioral functions.

However, several recent studies claim to have found new categories in rat USVs outside of the dichotomous framework along with evidence to infer their behavioral functions, *e.g.*, 30 kHz for responding to high-short USVs (5); 40 kHz for feeding (6); 31 kHz for social inequality and isolation (7). In light of these claims, revisiting past literature finds that there are indeed several reports on USVs whose acoustic features go beyond the dichotomous framework. For example, a previous study reported a possible cluster of USVs outside acoustic features of dichotomous USVs (8). Given these findings, we speculated that the practice of analyzing USVs by classifying them into 50- or 22-kHz categories based on a single boundary might have undermined the possibility of detecting new vocalizations and corresponding behavioral contexts. Therefore, we systematically surveyed publications on rat USVs (***Fig. 1C***) to provide evidence supporting the existence of vocalizations outside the two conventional categories. We also present a perspective for future studies on the behavioral functions of USVs beyond the dichotomy.

Results and Discussion

We found 413 articles based on an online database search and included 378 original research papers for adult rats' USVs that were available in English. We then comprehensively examined the description of USV features reported there and eventually extracted the frequency and duration of USVs from 84 articles (***SI Dataset S1***), excluding data obtained from physiological/genetic manipulations.

Firstly, we assessed how these data cluster on a duration-frequency plane by fitting 2-D Gaussian distributions. As a result, we found that data distributed as a mixture of two components corresponding to the dichotomous USVs (high-short and low-long) and a component that contained any other USVs (***Fig. 2A***). The centers of the three components ($C_{1\text{-}3}$) resembled the representative frequencies of dichotomous USVs, and their intermediate values (C_1: 56.95 kHz, 31.25 ms; C_2: 23.78 kHz, 908.15 ms; C_3: 43.49 kHz, 49.68 ms). Additionally, automatic clustering methods consistently showed clear clusters around the two dichotomous USVs. Although dichotomous USVs exhibit well-clustered distributions, while 41 cases in a total of 257 data points fall outside

the range of dichotomous USVs, which we named "non-dichotomous" USVs (***Fig. 2B***). Such cases were found in 19 articles (***SI Dataset S2***). The distribution of duration-frequency data points exhibits sufficiently different clusters for dichotomous and non-dichotomous USVs, *i.e.,* non-dichotomous USVs are unlikely artifacts due to inappropriate preprocessing.

Our literature survey revealed that non-dichotomous USVs have been reported in a non-negligible number, which supports recent studies that claimed new categories of USVs outside the 50-/22-kHz dichotomous framework. Since various types of non-dichotomous USVs have been repeatedly reported in both sexes and different strains (***Fig. 2C***), it is safe to assume that they are not exceptional or outliers nor biased toward any particular strain or sex. Although causal manipulations of behavioral contexts are required to establish functional categories, the results indicate that there may be far more acoustic categories of USVs than previously thought.

Our results do not negate the existence of traditional high-short and low-long USVs but suggest the existence of diverse USVs beyond the dichotomous framework. This raises the possibility that rat USVs can be mapped onto a continuum, not just a few categories. In line with this idea, recent studies revealed that laboratory mouse USVs do not show clear clusters but form a continuum (9). Thus, it may be reasonable to assume that the continuous acoustic structure of USVs is linked to the various emotional states. In conclusion, our study supports the abundance of non-dichotomous USVs beyond the 50- / 22-kHz dichotomous framework, strongly encouraging research on diverse vocalizations. To avoid erroneous categorization of USVs in future studies, we suggest not using numerical labels from now on. Future studies investigating the connections between vocalizations and emotions from the non-dichotomous perspective will provide deeper insights into the mechanisms of emotional communication in animals.

Materials and Methods

Data collection. We searched for rat's USV-related original papers in the literature database ("Web of Science"). We collected numerical descriptions of USV data (duration, frequency, sample size, sex, and strain) in papers. We carefully excluded data from pups, electric or optogenetic stimulation, and genetically or pharmacologically manipulated animals, whereas wild-type and saline controls were included (see details in ***SI Extended Methods***).

Clustering and boundary estimation. To estimate the boundary between dichotomous and non-dichotomous ultrasonic vocalizations, we first evaluated how data points distribute on the duration-frequency plane (logarithmic axis), by approximating the data using a Gaussian mixture model (3 components). Then, we attempted to find clusters for dichotomous USVs (high-short and low-long) by four automatic clustering methods and calculate the percent agreement among them. From these fitting and clustering results, we proposed tentative boundaries for dichotomous USVs that approximately cover the clustering results (see details in ***SI Extended Methods***).

Acknowledgments

Authors thank W. Tecumseh Fitch for providing valuable comments on the earlier version of manuscript.

References

1. J. W. Anderson, The Production of Ultrasonic Sounds by Laboratory Rats and Other Mammals. *Science* **119**, 808–809 (1954).
2. B. Knutson, J. Burgdorf, J. Panksepp, Ultrasonic vocalizations as indices of affective states in rats. *Psychol Bull* **128**, 961–977 (2002).
3. J. M. Wright, J. C. Gourdon, P. B. S. Clarke, Identification of multiple call categories within the rich repertoire of adult rat 50-kHz ultrasonic vocalizations: effects of amphetamine and social context. *PSYCHOPHARMACOLOGY* **211**, 1–13 (2010).

4. S. M. Brudzynski, F. Bihari, D. Ociepa, X.-W. Fu, Analysis of 22 kHz ultrasonic vocalization in laboratory rats: Long and short calls. *Physiology & Behavior* **54**, 215–221 (1993).
5. A. C. Berz, M. Wöhr, R. K. W. Schwarting, Response Calls Evoked by Playback of Natural 50-kHz Ultrasonic Vocalizations in Rats. *Front. Behav. Neurosci.* **15**, 812142 (2022).
6. G. Champeil-Potokar, *et al.*, Rats chirp with their mouth full: During an experimental meal, adult male Wistar rats emitted flat ultrasonic vocalisations upon feeding. *Front. Behav. Neurosci.* **17**, 1089631 (2023).
7. S. Okabe, Y. Takayanagi, M. Yoshida, T. Onaka, Novel 31-kHz calls emitted by female Lewis rats during social isolation and social inequality conditions. *iScience* **26**, 106243 (2023).
8. N. Takahashi, M. Kashino, N. Hironaka, Structure of Rat Ultrasonic Vocalizations and Its Relevance to Behavior. *PLoS ONE* **5**, e14115 (2010).
9. J. Goffinet, S. Brudner, R. Mooney, J. Pearson, Low-dimensional learned feature spaces quantify individual and group differences in vocal repertoires. *eLife* **10**, e67855 (2021).

Figures

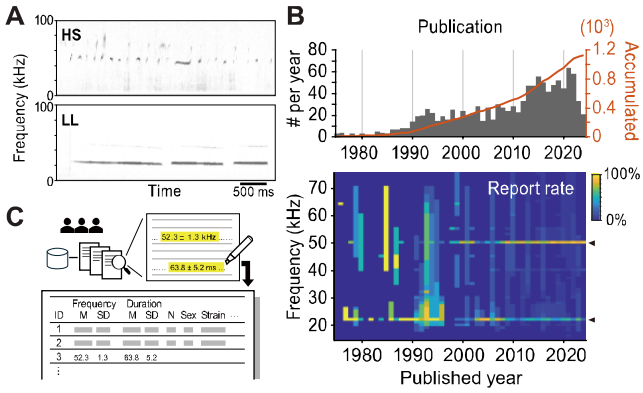

Figure 1. Recent publication on rat's USVs. (*A*) Typical spectrograms of “high-short” (HS) and “low-long” (LL) calls in rat’s USVs. (*B*) The number of publications per year on rat USV (upper) and appearance ratio of frequency values (kHz) in the abstracts of publications per year (lower). Reports of 50 and 22 kHz have been predominant (as arrowheads point). (*C*) Procedures of the paper survey and extraction of numerical values. Over 400 articles were selected from those published before 2024, and the numerical values of frequency, duration, and other information in the text were extracted.

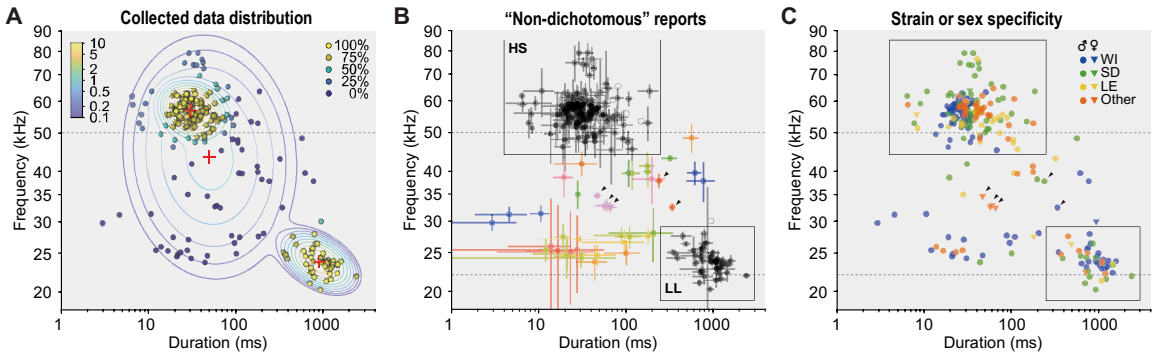

Figure 2. Distribution and clustering of collected data on a duration-frequency plane. (*A*) "Dichotomous"-likeliness of data and Gaussian mixture fitting. Filled circle shows representative frequency and duration reported in collected literature. Circle color indicates agreement rate (%) across four clustering methods for assigning the datapoint to the dichotomous call category. Contour line shows fitted Gaussian distributions (k = 3). Red cross indicates the center of distribution. (*B*) "Non-dichotomous" USVs found in collected data. Colors other than black correspond to the respective article reporting non-dichotomous USVs. Thick and thin error bars are SD and min-max ranges, respectively. Open circles indicate that representative values were reported without ranges. Square areas define the boundary for dichotomous USVs (HS and LL). Arrowhead indicates data from papers cited in the main text (5–7). (*C*) Sex and strain difference showed no specificity in the acoustic features (WI: Wistar, SD: Sprague-Dawley, LE: Long-Evans).

Supporting Information for

Beyond dichotomy: diversity of rat's ultrasonic vocalizations

Reo Wada, Shiomi Hakataya, Ryosuke O. Tachibana, Tomoyo Isoguchi Shiramatsu, Tetsufumi Ito, Kouta Kanno, Riseru Koshiishi, Jumpei Matsumoto, Yumi Saito, Genta Toya, Shota Okabe, Kazuo Okanoya

Extended Methods

Yearly publication analysis. To investigate the change over time in the number of papers published that mention rat vocalization, we searched the following search key in the Pubmed paper database: "((vocalization[Title/Abstract]) OR (vocalizations[Title/Abstract])) AND (ultrasonic[Title/Abstract]) AND ((rat[Title/Abstract]) OR (rats[Title/Abstract]))". As a result, 1134 papers were found (searched on August 9, 2024). The number of published papers was counted per year.

Call frequency reporting rates in paper abstracts. We searched for USV-related papers the following search key in the Web of Science: "AB=(kHz (rat OR rats) (vocalization OR call OR vocalizations OR calls))". As a result, 567 papers were found (searched on August 13, 2024). Descriptions of frequency in kHz in abstracts of the papers were automatically detected as single point numbers or range expressions, such as "X-Y kHz", "between X and Y kHz", etc. Unclear range descriptions (i.e., shown with "above", "higher than", ">", "<", etc.) were excluded from further analysis since almost all of them were related to analysis or stimulus conditions. After these extractions, we counted the detection for every frequency number from 15 to 80 kHz in 1 kHz steps. When the same frequency was reported multiple times in a paper, we counted it only once to avoid redundancy. For the range expressions, we added one for all frequency bins within that range. We then created a two-dimensional histogram with publication year as the horizontal axis and frequency as the vertical axis. The raw histogram was divided by the number of papers that reported the frequency in their abstracts to obtain the frequency reporting rates for each publication year.

Paper survey and selection for acoustic parameters. We searched for USV-related papers using the following search key in the Web of Science: "AB=(rat (ultrasonic OR kHz) (call OR vocalization) NOT(pup OR infant OR neonatal))", Document type: "Article", Citation Topics Micro: "1.5.1511 Ultrasonic Vocalization". We only included papers that were classified in the following "WoS Categories": Neurosciences, Behavioral Sciences, Pharmacology Pharmacy, Psychiatry, Psychology Biological, Zoology, Multidisciplinary Sciences, Clinical Neurology, Veterinary Sciences, Psychology, Psychology Multidisciplinary, Biology, Audiology Speech Language Pathology, Developmental Biology, Endocrinology Metabolism, Substance Abuse, Medicine Research Experimental, Acoustics, Genetics Heredity, Physiology, Anesthesiology, Linguistics, Cell Biology, Ecology, Engineering Biomedical, Environmental Sciences, Evolutionary Biology, Otorhinolaryngology, Psychology Experimental. As a result, we found 398 articles (searched on April 18, 2024). We further added 15 papers that authors found by chance during the article survey, and hence, the literature list size became 413. After collecting these articles, we started to read the literature and omitted articles that were reviews or book chapters but not original papers.

To extract the USV data (duration, frequency, sample size, sex, and strain), we marked the corresponding descriptions in paper PDFs by drawing highlights using the reference management software Zotero. We carefully excluded USV data obtained from pups, electric or optogenetic stimulation, and genetically or pharmacologically manipulated animals, whereas wild-type and saline controls were included.

Many papers in our collection described the "mean frequency" or "peak frequency" as the frequency characteristic of targeted USV syllables. The mean frequency typically indicates the mean of the frequency trace in a syllable, whereas the peak frequency is the frequency at the time point where the amplitude showed its peak in the syllable. In rare cases, we found different

descriptions, such as "representative frequency" or "dominant frequency." We labeled these characteristics differently during the data extraction. The USV characteristics were typically described as "mean ± SD" or "mean ± SE" in most articles that were calculated across subjected animals. In some cases, we also found median values or the value range as "minimum - maximum" forms. We marked these statistics differently.

For simple visualization in the main figure (***Fig. 2***), we dismissed the differences in the frequency characteristics (*i.e.*, mean, peak, or other). We converted the standard error (SE) to standard deviation (SD) if we could obtain the corresponding sample size (N) by multiplying the SE value by the square root of N.

Estimation of dichotomous vs non-dichotomous boundary. In our meta-analysis, we tentatively determined the boundary between dichotomous and non-dichotomous USVs (referred to as the "D-ND boundary"). The duration-frequency data points appeared to form sufficiently distinct clusters for dichotomous and non-dichotomous USVs, suggesting that non-dichotomous USVs are not artifacts arising from improper preprocessing. However, the exact location of the D-ND boundary was not consistently identified across various clustering methods. Thus, we propose tentative boundaries in this study that encompass the approximate results of multiple clustering analyses. The analysis details are as follows.

First, we evaluated the distribution shape of the scatter plot of the duration-frequency data. By approximating the data using a Gaussian mixture model (#components: k = 3) on the logarithmic axis, we found that the distribution was a mixture of two components corresponding to dichotomous USVs and a component that contained other USVs. The centers of the three components resembled the representative frequencies of dichotomous USVs and their intermediate values, suggesting that dichotomous USVs show well-clustered distributions and non-dichotomous data points are unlikely to belong to the dichotomous distributions.

Next, we categorized data points into three groups (high-short, low-long, and other calls) using four automated clustering methods:

1. Clustering based on the previously obtained Gaussian mixture approximation.
2. DBSCAN (1, 2), which identifies clusters as high-density areas separated by areas of low density regions (parameter setting: *eps* = 0.4, *min_samples* = 20).
3. OPTICS (3), similar to DBSCAN but a relaxed constraints (*min_samples* = 20, *max_eps* = 0.5).
4. BIRCH (4), which approximates a hierarchical tree-like structure of subclusters (*threshold* = 0.5, *branching_factor* = 50, *n_clusters* = 4).

We standardized the frequency and duration values by dividing them by their standard deviation after logarithmic conversion. For the first three methods, parameters were optimized to clearly visualize clusters corresponding to dichotomous USVs while dividing data into three clusters. For BIRCH, the maximum number of clusters was set to four. Clustering was performed using functions available in MATLAB Statistics and Machine Learning Toolbox (R2024b) and Python scikit-learn library (version 1.5.2, Python version 3.12).

All four clustering methods consistently detected two core regions corresponding to high-short and low-long calls. Additionally, each method identified a third cluster separate from these core regions (classified as outliers by DBSCAN and OPTICS, and as two combined clusters by BIRCH). This result supports the conclusion that non-dichotomous USVs are genuine and not simply artifacts removed from the tails of dichotomous distributions during preprocessing. However, examining marginal data points closely revealed inconsistencies across methods regarding cluster assignments. Such discrepancies likely stem from the relatively small dataset size rather than an absence of an actual boundary.

Although a similar tendency was observed across the clustering methods, the results remained inconclusive regarding the precise determination of D-ND boundaries. Therefore, considering these results, we conservatively defined tentative boundaries to include as many data points as possible in the dichotomous USV range. Rat USVs typically exhibit 'trills'—intense frequency modulations with repetitive up-down movements—under positive conditions (50-kHz USVs). These trills elevate the syllable-average frequency by approximately 10–30 kHz, suggesting that data points around 70–80 kHz likely represent trill-containing syllables. Consequently, frequency boundaries were set to encompass these points within the dichotomous

USVs. Additionally, syllable durations can vary based on segmentation criteria and allowed gap lengths; thus, we included flexibility in our duration boundaries. The clusters corresponding to the 22-kHz USVs demonstrated sufficient cohesion, prompting us to establish symmetrical boundaries around the distribution center. Based on these considerations, we tentatively set the D-ND boundaries as follows: the high-short region (50-kHz USV) at 44–85 kHz and 4–250 ms, and the low-long region (22-kHz USV) at 19–29 kHz and 250–3000 ms.

Dataset S1 (separate file). Duration, frequency and other data collected from 84 papers.

Dataset S2 (separate file). Duration, frequency and other data of non-dichotomous USVs (41 datapoints, 19 papers).

SI References

1. M. Ester, H.-P. Kriegel, J. Sander, X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise. *kdd* **96**, 226–231 (1996).
2. E. Schubert, J. Sander, M. Ester, H. P. Kriegel, X. Xu, DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN. *ACM Trans. Database Syst.* **42**, 19:1-19:21 (2017).
3. M. Ankerst, M. M. Breunig, H.-P. Kriegel, J. Sander, OPTICS: ordering points to identify the clustering structure. *SIGMOD Rec.* **28**, 49–60 (1999).
4. T. Zhang, R. Ramakrishnan, M. Livny, BIRCH: an efficient data clustering method for very large databases in *Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data - SIGMOD '96*, (ACM Press, 1996), pp. 103–114.