Quantifying Sex Differences in Behavior in the Era of "Big" Data

Brian C. Trainor[1] and Annegret L. Falkner[2]

[1]Department of Psychology, University of California, Davis, California 95616, USA

[2]Princeton Neuroscience Institute, Princeton, New Jersey 08540, USA

Correspondence: bctrainor@ucdavis.edu

Sex differences are commonly observed in behaviors that are closely linked to adaptive function, but sex differences can also be observed in behavioral "building blocks" such as locomotor activity and reward processing. Modern neuroscientific inquiry, in pursuit of generalizable principles of functioning across sexes, has often ignored these more subtle sex differences in behavioral building blocks that may result from differences in these behavioral building blocks. A frequent assumption is that there is a default (often male) way to perform a behavior. This approach misses fundamental drivers of individual variability within and between sexes. Incomplete behavioral descriptions of both sexes can lead to an over-reliance on reduced "single-variable" readouts of complex behaviors, the design of which may be based on male-biased samples. Here, we advocate that the incorporation of new machine-learning tools for collecting and analyzing multimodal "big behavior" data allows for a more holistic and richer approach to the quantification of behavior in both sexes. These new tools make behavioral description more robust and replicable across laboratories and species, and may open up new lines of neuroscientific inquiry by facilitating the discovery of novel behavioral states. Having more accurate measures of behavioral diversity in males and females could serve as a hypothesis generator for where and when we should look in the brain for meaningful neural differences.

SEX DIFFERENCES IN BEHAVIOR MAY BE SIMULTANEOUSLY OVERLOOKED AND OVEREMPHASIZED

Early in the study of animal behavior, dramatic sex differences observed in courtship and mate selection behaviors launched the idea that behaviors themselves were under evolutionary control, and enforced long-held assumptions that there might be "male" behaviors and "female" behaviors (Darwin 1882). We now know that sex differences in investment in reproductive and mating strategies can lead to clear and quantifiable sex differences in behavior that directly affect the fitness of an individual. Historically, the study of sex differences focused on the most extreme differences in behavior including reproductive behaviors (Ball et al. 2014) and aggression (Lischinsky and Lin 2020). However, the absence of an obvious link to parental or reproductive strategy does not mean that there are not important

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a039164

1

B.C. Trainor and A.L. Falkner

sex differences in behavior, let alone in the underlying neural dynamics or circuit architecture that generate the behavior (De Vries and Boyle 1998). Although it may be obvious to look for sex differences in social behaviors that are directly related to mating and aggression, there is growing evidence that subtle sex differences may have been systematically overlooked.

In contrast, although some sex differences in behavior may have been ignored through a hyperfocus on behaviors with direct links to reproduction, described sex differences in behaviors that "do" have direct links to reproductive strategy may, in some cases, be overblown. For example, in many species males but not females use vocalizations to attract mates or compete with other individuals. There has been a recent appreciation that in some species females also use vocalizations in competitive interactions (Kelly 1993; Emerson and Boyd 1999). Thus, descriptions of some behaviors as sexually dimorphic may be reinforced by existing biases about "male" and "female" behaviors, and deserve further scrutiny.

Implementing more complete descriptions of behavior for both sexes is an important remedy for overcoming these biases. We frequently lack complete behavioral descriptions for both sexes as, for many years, females were not included in behavioral analyses due either to neglect or misguided beliefs. Females were often excluded from studies because of the concern that hormone fluctuations across the reproductive cycle would obscure the neuroscientific objectives. This rationale has been discredited (Prendergast et al. 2014; Becker et al. 2016) and the consideration of "sex as a biological variable" (SABV) by the NIH has been a critical and necessary corrective to a decades-long imbalance in neuroscience (but see Mamlouk et al. 2020). When this more balanced approach is applied to more foundational behavioral "building blocks" such as reward learning or defensive behavior, subtle sex differences in these behaviors are sometimes found hiding in plain sight.

One example of a "missed" behavior is the recent finding that female rats may use a different behavioral strategy than male rats during the expression of learned fear. Historically, learned fear has been quantified by observing canonical freezing behavior. In many fear conditioning studies females show reduced freezing relative to their male counterparts (Fig. 1A; Maren et al. 1994; Pryce et al. 1999; Clark et al. 2019). However, through careful analysis of video recordings, researchers identified a new behavior—"darting"—that was far more prevalent in females than in males (Fig. 1B; Gruene et al. 2015).

Why does this matter? If a field, as a result of several decades of exclusive research on the behavior of males, has been focused on freezing as the de facto defensive reaction, then the responses of the females, who freeze less often, appear aberrant. The fact that females use more diverse strategies to defend against threat should prompt a rethink about what it means when an animal freezes (or does not freeze) and how this behavioral diversity should be taken into account when assessing how animals learn. The identification of darting behavior was in some ways accidental. It was the product of taking a wider view of which behaviors individuals of both sexes performed during a routine experimental paradigm: fear conditioning. The accidental nature of this discovery raises the question of whether a more systematic approach can be used to evaluate whether there are important sex differences in other behaviors. As we enter a new era in the ability to collect and analyze behavior data with more sophistication and quantitative rigor, it may be time to ask how we can best harness new technology to discover subtle, but potentially important, patterns in behavior and reduce our reliance on single-variable readouts for internal states.

Several recent reviews have called for a return to a more ethologically motivated study of behavior in neuroscience (Anderson and Perona 2014; Krakauer et al. 2017; Datta et al. 2019). Implicit in these arguments is that there is new knowledge to be gained by moving the lens away from the brain and instead becoming more focused on behavior, and that insights about behavior can be used to direct neuroscientific inquiry. However, many of the examples used in these calls to action do not explicitly consider behavior in males and females. To take a more holistic look at behavior, individual differences
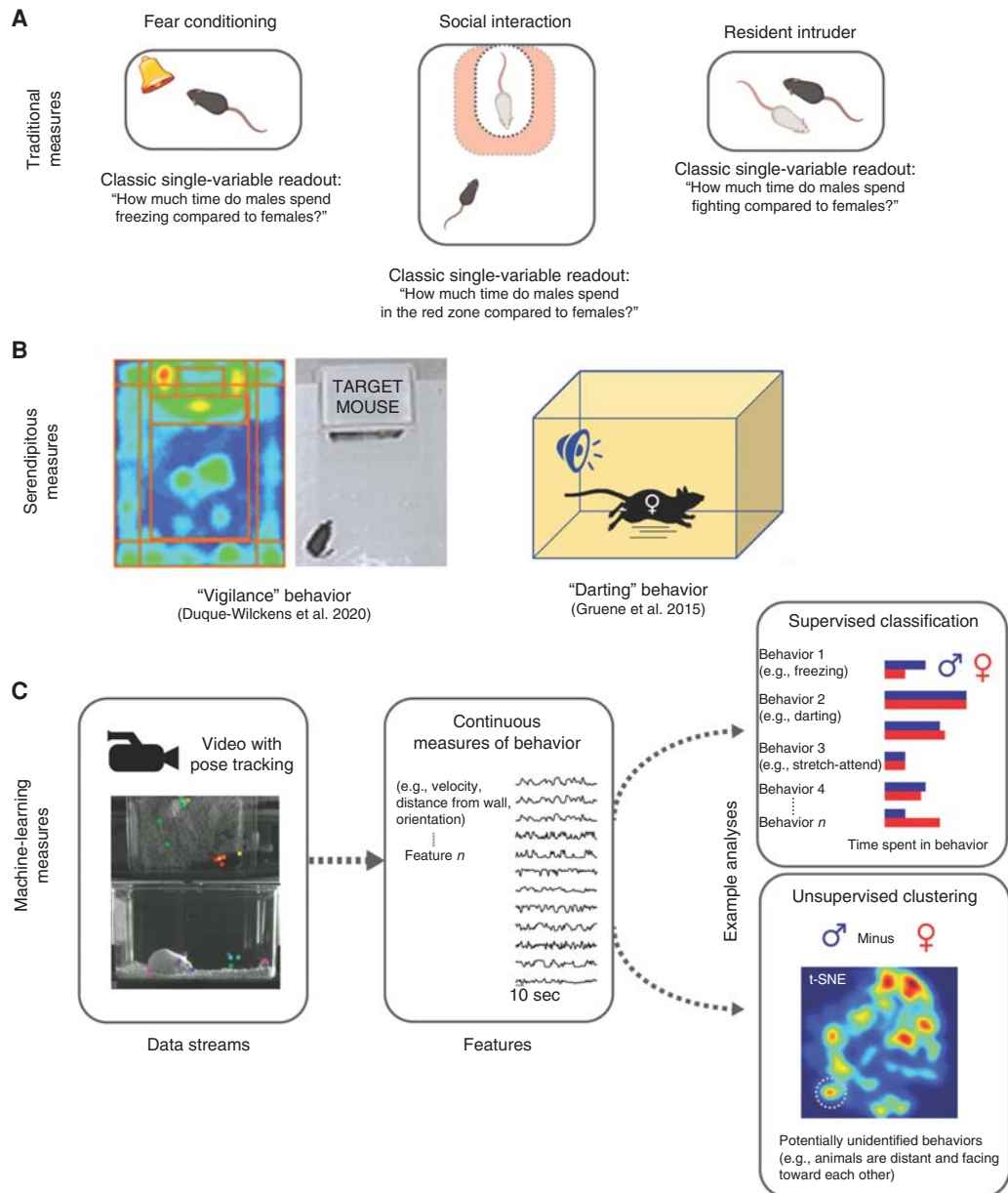
Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a039164

Figure 1. New tools to facilitate the discovery of sex differences in behavior. (*A*) Common behavioral paradigms used in neuroscience commonly rely on single-variable descriptions of complex behaviors for both sexes. Examples include fear conditioning (*left*), social avoidance (*center*), and the "resident intruder" paradigm for social behavior (*right*). (*B*) Recent examples of serendipitously discovered behaviors in these paradigms suggest that single variables are insufficient to describe sex differences in behavior. Examples include increased social vigilance in females (*left*), and increased "darting" behaviors in females during fear conditioning (*right*). (*C*) Adopting machine-learning strategies for pose tracking and video analysis will facilitate behavioral discovery by allowing the robust quantification of multiple behaviors, and may enable the discovery of new behaviors. One example is using video acquisition and pose tracking (in this case, DeepLabCut), to reveal behavioral "features" from supervised analyses (*right top*), but also applying unsupervised analyses (*right bottom*) to provide a more nuanced view of sex differences in complex behavior. Difference heat map may reveal behavior clusters that are enriched in one sex over the other.

B.C. Trainor and A.L. Falkner

across both sexes must be measured, such that appropriate behavioral readouts that fully represent the behavior of both sexes can be designed.

Single-variable measurements for complex internal states abound in the scientific literature. Initial developments in the computerization of behavioral analysis automated a narrow set of measurements allowing high-throughput behavioral analyses. However, simply measuring time in the center of an open field, or time spent in an interaction zone during a social interaction test, can overlook other important behavioral patterns (von Ziegler et al. 2021). Developments in machine vision and machine learning provide an opportunity to apply neuroethological approaches without losing the throughput of automated behavioral scoring. Often behavioral analysis is used to infer an underlying state such as fear or motivation. Using more complete descriptions of behavior in both sexes under ethologically relevant conditions is likely to provide richer insights into these states. The reasons for this are simple: Animals of different sizes, physiology, or life experience might apply different strategies for deciding which behaviors to use and when. Using the same logic for why we might use age-matched cohorts to study a particular behavior (because age-matched cohorts should be of comparable size and experience), we can extend this same logic to looking at sex differences in a broad range of behaviors. The use of machine learning confers additional advantages, since this approach may enable the detection of behaviors or behavioral sequences that would be difficult or impossible for humans to detect. Historically, behaviors have been classified by experts with deep knowledge of the species being studied. However, even an expert could be biased by the way the human sensory system perceives motor patterns. The development of novel algorithms that detect or classify behavior with minimal or no human intervention could provide more unbiased descriptions of behavior and potentially even identify previously undescribed behaviors missed by experts.

Here, we will survey a brief history of general approaches to quantifying sex differences in behavior, and consider new directions for the field as it incorporates new computational tools

to simplify increasingly high-dimensional and multimodal data. In addition, we will propose a model of how researchers might use this information to guide neural circuit interrogation of how these behavioral differences are enforced.

QUANTIFYING SEX DIFFERENCES WITHOUT BIAS

Then and Now: A Brief History of Quantifying Behavioral Difference

Studying sex differences in behavior has deep roots in neuroethology, because some of the first well-described behaviors, including egg-protecting maneuvers famously described by Karl Lorenz, are sexually dimorphic. These discoveries were based on human observations performed either directly or through recordings. Several methodologies were designed to account for the limits of human abilities (Martin and Bateson 1994). To obtain more detailed observations, an investigator could use "ad libitum" sampling and note down whatever is visible or relevant during an observation period. This approach provides flexibility to record unusual behaviors, but is prone to overlooking subtle responses, especially in a group setting (Bernstein 1991).

An alternate approach is scan sampling, in which individuals are observed at brief, regular intervals and the observer uses a defined list of behaviors to record. Scan sampling allows for the monitoring of multiple individuals but at limited depth. Furthermore, the brief scans will usually miss infrequent behaviors. Focal sampling, in which the investigator focuses observations on a single individual, is more conducible to detecting rare behaviors. Similar to scan sampling, the behaviors to be scored are decided in advance, and in laboratory settings the focal animal is usually visible for the entire observation period. If combined with video recordings, it is possible for an investigator to intensively observe behavior from multiple individuals interacting at the same time. However, this approach is laborious, as the investigator must score the same recording multiple times, each time focusing on a different individual. Although it is usually possible for multiple observers to be trained to score behav-

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a039164

ior with high reliability, there is a limit to what can be realistically expected from human observers. In addition, some interactions between individuals may occur synchronously and may not be readily identified by observers focused on individual actions rather than those of a group or pair. Thus, even heroic efforts on the part experimenters using existing methods are likely to miss important behavioral sex differences that could potentially provide significant insights into observed neural differences. The adoption of semi-automated and automated methods for behavioral capture, including behavioral "pose" tracking, could add a new level of rigor to behavioral analyses that can identify new behavior patterns or overlooked sex differences.

Semi-Automated and Automated Methods

One of the first forms of automated behavioral tracking was computerized tracking of beam breaking as a measure of general locomotor "activity." This allows monitoring of more individuals over a longer period of time. One example of a sex difference in locomotor behavior is seen in response to pharmacological compounds. Female rodents show an exaggerated locomotor response to cocaine compared to males (Roth and Carroll 2004; Van Swearingen et al. 2013), a difference that appears to be established before birth (Forgie and Stewart 1993). The development of commercial video tracking systems led to more widespread use of computerized analysis of behavior as it allows more flexibility in the types of apparatuses that can be used and for more parameters to be quantified. One useful variable is path length, which is laborious to calculate by eye and provides a more accurate measure of search strategy in learning and memory tests such as the water maze. For example, although male rats took shorter paths to reach the hidden platform than females, this difference was eliminated if rats were given an initial training session to familiarize them with the requirements of the water maze tests (Perrot-Sinal et al. 1996). One drawback to video tracking methods is that subtle or complex behaviors can be difficult to detect when data collection is limited to x, y coordinates.

Although human-driven behavioral annotation of more subtle or complex behaviors remains a staple of behavioral analysis, the ability to speed up this process has been assisted by several pieces of software that facilitate video "tagging," including open source options like the Behavioral Observation Research Interactive software (Friard and Gamba 2016). These software suites enable users to examine video frame-by-frame and "tag" relevant behavioral moments (e.g., the onsets and offsets of user-defined behaviors). However, for aggressive behaviors, there is significant variability in how individuals score behaviors (Segalin et al. 2020). Furthermore, this approach is time consuming and relies on behavior patterns defined by the experimenter, which themselves may be biased. Thus, these methods may be less conducive for discovering novel behaviors or sex differences.

A significant advance in our ability to quantify complex behavior arrived in the form of pose tracking, in which specific body parts are tracked to define behavioral patterns. Although tracking the pose of individuals is not a new idea (Marr and Nishihara 1978), the ability to do it reliably, inexpensively, and with limited computational knowledge, is novel. Motivated by significant recent advances in automated video analysis, including the use of trained "deep" convolutional neural networks to reliably identify the content of an image, these methods have revolutionized the ability of researchers to study animal behavior in the laboratory and beyond (Graving et al. 2019; Li et al. 2020). In particular, machine-learning-based methods, including user-friendly software such as DeepPoseKit, LEAP, Deep-LabCut, or MARS, allow users to perform automated detection of user-defined points (e.g., the forelimb or nose of a mouse) across frames of a given video data set (Mathis et al. 2018; Graving et al. 2019; Pereira et al. 2019; Segalin et al. 2020). These algorithms rely on a user-generated training data set, in which the user labels relevant body parts (e.g., an animal's forelimb or nose) in a subset of videos to train the algorithm. After training, the algorithm can then detect the presence or absence of those selected features in the remaining data set. Depending on the video quality and the discriminability of the features

B.C. Trainor and A.L. Falkner

being tracked, pose tracking can be implemented with little or no algorithmic "proofreading," the process of iteratively refining the model. Importantly, the use of these methods is scalable. With previous methods, manually tracking multiple features or even multiple animals significantly increases the time cost to the experimenter. However, the use of automated pose-tracking approaches allows the application of large numbers of labels to an ever-expanding video data set. An important point to note is that pose tracking, which ultimately results in a time series of spatial locations of desired body parts, may not directly yield usable behavioral data. Instead, these spatial locations are often first converted into relevant features that represent specific relationships between individual points (e.g., the nose-to-nose distance between two interacting rodents). To extract additional meaningful information from pose-tracking data, either supervised or unsupervised classification strategies based on the fields of statistics and machine learning can be applied. Several excellent recent reviews thoroughly outline how these methods may be used to quantify behavioral dynamics (Mathis et al. 2020; Pereira et al. 2020), but we will summarize their use briefly here.

Supervised strategies, including behavioral classification, are frequently used to identify instances of a behavior in a larger data set. These strategies rely on a user-generated ground truth, in which an expert labels instances of a user-specified behavior (e.g., are mice fighting in this frame?) to train the algorithm to link specific poses with these behaviors on a frame-by-frame basis. Recent examples of this method used support vector machines and random forest classification to assist in identifying instances of user-defined behaviors from pose-tracking data (Nilsson et al. 2020). A strength of this method is that, because the user defines the behaviors to be tracked, the end data are straightforward to interpret. For example, the frequency, bout length, or duration of a particular behavior can be compared in two groups. However, this strength is also a constraint: detected behaviors are limited to what the user has already specified as being important, which precludes the ability to identify novel behavioral patterns. For exam-

ple, some sex differences in behavior may lie outside our expectations (e.g., darting behavior). A second drawback to these strategies is that they perform poorly in situations in which data are limited or behaviors of interest are extremely rare or ambiguously defined by the user. Therefore, additional methods may need to be used that allow for the identification of "new" behaviors.

Unsupervised methods, which infer the structure of the data without ground truth labeled outputs will likely yield additional insights. Although supervised methods excel at extracting patterns to assist in identifying the "known unknowns," unsupervised methods can be used in parallel to identify behaviors that might have been missed by experimenters (the "unknown unknowns"), or to identify latent structure in the data. Successful examples of this approach are beginning to emerge using pose-tracking data as the input. For example, using a nonlinear embedding approach on pose-tracking data from freely moving individuals revealed interpretable behavioral "clusters" (Berman et al. 2014; Klibaite et al. 2017), in which each cluster represented an identified behavior (e.g., "turning left" or "grooming"). Other clusters may represent behavioral states (e.g., two individuals interacting socially are at a specific distance and orientation relative to each other) that may not, to the human observer, initially appear to be distinct "behaviors." However, further analysis could prove that motor patterns identified by unsupervised analyses may, in fact, be informative about the internal state of the animals or how they pattern their next choice of behavior. Open source tool kits, including B-SOiD, have made unsupervised data exploration more tractable for the end user (Hsu and Yttri 2021). In head-fixed preparations for imaging and electrophysiology, high-speed video of the face and body is often collected alongside neural data. Significant progress has been made using supervised and unsupervised approaches to detect signatures of facial expression or emotional state (Khan et al. 2020; Dolensek et al. 2020), which previously could only be identified with painstaking frame-by-frame analysis by human observers (Grill and Norgren 1978). However, even with these automated methods,

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a039164

sex differences are rarely reported. Given the strong evidence for sex differences in physiological and behavioral responses to stress (Laman-Maharg and Trainor 2017), it seems likely that males and females could differ in the use of facial expressions.

Although these tools provide new ways to discover new behaviors, scientists with strong training in behavioral analyses are still needed to assess the importance of motor patterns identified by unsupervised methods and design experiments to determine their function.

INCORPORATING MULTIMODAL DATA STREAMS

Of course, not all behavior is detectable bodily movement. Other multimodal data streams can be integrated with pose tracking to create a more holistic picture of behavior. Some forms of communication, for example, are nearly impossible to detect through limb tracking (although they might be related). For many rodent researchers, recording ultrasonic vocalizations (USVs) has revealed critical insights about courtship, parental, and prosocial behaviors (Holy and Guo 2005). In addition to auditory streams of data, other important behavioral variables, including changes in body temperature (using temperature-sensitive cameras) or odor profile (using a photoionization detector) can be integrated with ongoing movement detection to provide additional behavioral context.

In addition to their use in video data, supervised and unsupervised strategies have also been successfully applied to the detection and segmentation of "big" audio data, in particular to the identification of rodent calls (Van Segbroeck et al. 2017; Coffey et al. 2019; Vogel et al. 2019). For communication signals emitted in groups of animals, experimenters have to solve two specific problems: (1) What is the call being emitted (call classification), and (2) Who is making this call (call identification)? These problems pose unique challenges because audio data is frequently collected from multiple spatially separated microphones, and because calls are made during social interactions, communication signals are often overlapping from multiple individ-

uals, making it difficult to determine which individual is vocalizing.

For call classification, using a combination of manual call identification, supervised classified calls, and unsupervised classified calls, the full repertoire of rodent communication signals is starting to emerge. Although some algorithms for USV segmentation (e.g., DeepSqueak and MUPET) use the acoustic features of the audio signal to classify or segment the sound, use of variational autoencoders to define a latent space (a lower dimensional representation of the highly complex acoustic spectrogram) can be used to directly compare these models (Goffinet et al. 2021). These models have converged on a common set of vocal syllables that make up the full mouse vocal repertoire, which allow the content of calls to be compared between males and females. The ability to classify the content of calls can now be combined with positional information or pose-tracking data captured by video to allow recorded calls to be "assigned" to an individual. These data, combined with behavioral analysis described above, further allows additional insights about which behaviors are correlated with specific vocalizations.

Vocal communication in rodents and birds has long been believed to be strongly sexually dimorphic, with females vocalizing rarely, if ever. However, use of more unbiased methods for audio analysis combined with positional information has revealed that female mice, rather than being silent recipients of male communicative signals, are active participants (Neunuebel et al. 2015), in some cases contributing nearly 20% of the total vocalizations. A detailed analysis of the structure of female calls shows that they are largely similar to the calls of the male (Hammerschmidt et al. 2012), but are more likely to call to other females than to males. This second look at the prevalence of female song in rodents is echoed in recent work in birds, which suggests that birdsong in many species is not only prevalent in females, it is conserved across species (Odom et al. 2014). This finding represents a good example of how use of an unbiased behavioral lens to look at sex differences decreases the size of an assumed sex difference.

B.C. Trainor and A.L. Falkner

AUTOMATED BEHAVIORAL PHENOTYPING TO ASSESS INTERNAL AND AFFECTIVE STATES

The use of machine-learning approaches to behavioral analysis can be used not only to assess differences in moment-to-moment behaviors, but also to make predictions about an individual's internal or affective state. Sex differences in these metrics of well-being have already been observed using automated and semiautomated methods. For example, acute behavioral phenotyping in the home cage shows sex differences in the stress responses of mice, in which females show more depression-like phenotypes than males such as reduced grooming and general activity (Goodwill et al. 2019). These effects were reversed by antidepressant treatment and echoed more traditional tests of rodent depression-related behavior such as sucrose anhedonia. In a more long-term experiment, full-time "24/7" monitoring of mice of both sexes in their home cages across long timescales reveals sex differences in how individuals respond to a cage change stress, with males showing a more variable stress response to cage changing (Pernold et al. 2019). These methods can also be used to further quantify individual variability both between and within sexes. Such "behavioral phenotyping" has been recently used in fish and mice to quantify individual differences in behavior in response to various pharmacological perturbations (Hoffman et al. 2016; Wiltschko et al. 2020) and can be further extended to look at sex differences.

Understanding sex differences in the way individuals respond to stress is critical for developing behavioral methods to evaluate affective state. In several species of rodents, both males and females show strong approach responses to an unfamiliar social context (Duque-Wilckens et al. 2016; Yohn et al. 2019). The most widely used method to measure this response is a social interaction test (Fig. 1A) in which an unfamiliar target mouse is placed into a small wire cage and the time the focal mouse spends near this cage is quantified (Golden et al. 2011). This single approach to capturing the stress response is likely insufficient to capture behavioral variability across sexes. Exposure to social stressors (in the form of aggressive interactions) can induce different behavioral phenotypes in males and females. In California mice (*Peromyscus californicus*), three brief episodes of social stress exposure reduces social approach in females but not males (Trainor et al. 2011, 2013). If a large arena is used for testing, additional behaviors can be observed. When there is more space, stressed females orient toward the target mouse while simultaneously avoiding it, a response referred to as social vigilance (Duque-Wilckens et al. 2018). Social vigilance appears to function as a risk-assessment behavior in adverse or changing social environments (Wright et al. 2020).

Like darting behavior, social vigilance was largely discovered by accident. Heat maps produced from social interaction tests indicated that many stressed females would spend large amounts of time in the center of the arena (Fig. 1B; Greenberg et al. 2014). Initially this finding was baffling, because anxiety-like states are associated with decreased time spent in the center of an open field. Only after measuring several different variables produced by computer tracking software was it discovered that stressed females were orienting to a stimulus mouse while simultaneously avoiding it (Fig. 1B). In California mice, social stress exposure induces an enduring increase in social vigilance in females but not males 2 weeks later. The neural circuitry for social vigilance is intact in male California mice, as it is observed immediately after an episode of defeat. However, this effect is transitory (Duque-Wilckens et al. 2020). Social vigilance has also been observed in stress models using C57Bl/6J mice. In one approach, sexually experienced CFW females reliably showed aggression toward female C57Bl/6J in a resident-intruder test (Newman et al. 2019). C57/Bl6J females exposed to 10 such episodes of defeat showed reduced social contact in the home cage and increased social vigilance in an unfamiliar arena. In the unfamiliar arena, inclusion of social vigilance was important because standard measures of time spent in an interaction zone did not differ between control and stressed females. In an alternative approach, male or female C57Bl/6J observes another male experience social defeat (Warren et al. 2020). This vicarious exposure to defeat stress induces social

vigilance in both male and female C57Bl/6J, but the effect is significantly stronger in females (Duque-Wilckens et al. 2020).

This serendipitous discovery of social vigilance suggests that more holistic, quantitative descriptions of behavior could generate significantly more insight than the standard single-variable measure of affective state or stress response (e.g., measures of time spent in an interaction zone). Machine-learning approaches could open new doors for identifying important sex differences in behavioral responses to stress. Because paradigms including chronic social defeat and social interaction tests have been widely applied, standardization across laboratories in how behaviors are quantified and scored could generate new opportunities for generating large data sets that could be directly compared. For example, if video recordings were standardized, pose tracking could be used in a consistent way across laboratories and fed into either supervised or unsupervised algorithms to identify specific behavioral patterns that differ between sexes (Fig. 1C). Although social approach and vigilance can be quantified using comparatively simple video tracking software that tracks both the head and the base of the tail, other behaviors may be subtle. For example, rodents respond to both predators (Hubbard et al. 2004) and social threats (McCann and Huhman 2012) with a "stretch-attend" posture in which the individual crouches to reduce visibility and slowly approaches the threat. Stretch-attend is traditionally scored by trained observers and is more easily quantified from a side view versus the overhead view that is more typically used in social interaction tests. Deep-learning approaches that track multiple body points (Mathis et al. 2018; Pereira et al. 2019), combined with machine-learning tools (Sturman et al. 2020), are better suited to detecting these posture-specific behavior patterns than standardized video tracking systems.

MAPPING SEX DIFFERENCES IN BEHAVIOR TO SEX DIFFERENCES IN THE NERVOUS SYSTEM

Comprehensive, quantitative descriptions of behavior can generate hypothesis for the neural mechanisms that underlie putative differences. For example, sex differences in behavior can occur on a spectrum from completely binary in a single behavior, to extremely subtle and across a diverse set of behaviors. Knowledge about effect sizes across a range of behaviors may give clues to the neural mechanism underlying the difference.

When behaviors appear to have a bimodal distribution, that may indicate that one sex completely lacks the ability to perform this behavior, and clear sexual dimorphism might be observed in motor processing circuits. This is the case in some species of frogs in which males but not females use advertisement vocalizations to attract females and compete with other males. In African clawed frogs, the motor neurons and muscle fibers that control vocalizations are androgen-sensitive, and increased androgen levels during puberty are essential for normal development of the laryngeal muscles required for vocalizations (Tobias et al. 1993; Emerson and Boyd 1999). However, careful analysis revealed that in some species of frogs, females also produce advertisement calls (Kelly 1993; Emerson and Boyd 1999). The mechanisms for advertisement calls in females have been understudied (Wilczynski et al. 2017). A potentially important neuropeptide is arginine vasotocin, which can induce advertisement calls in female tree frogs even in the absence of androgen treatment (Penna et al. 1992). Although these calls have different frequencies than those of males, this example shows the importance of considering both sexes, even in cases in which behavior at first glance appear to be strongly sexually dimorphic.

However, for most sex differences in behavior there is continuous variability in frequency, intensity, or duration. One example is aggression. Despite the fact that in most species, males are more aggressive than females, this appears to be a matter of degree and not kind. Although males are in fact more aggressive in most species, female aggression is often displayed in different contexts. Instead, female aggression is most frequently displayed during lactation. Key brain areas for patterning aggressive motivation and action, including the ventromedial hypothalamus, ventrolateral area (VMHvl), and the periaqueductal gray (PAG), appear to be largely

B.C. Trainor and A.L. Falkner

similar in males and females (Lin et al. 2011; Hashikawa et al. 2017; Falkner et al. 2020). Rather than showing large structural differences, aggression circuits instead appear to differ in the number of neurons that are coactivated by same- and opposite-sex social partners. Although males have a "shared" population of VMHvl neurons that responds to both male and female conspecifics, VMHvl neurons in females that respond to males and females are functionally and anatomically separated (Falkner et al. 2014; Hashikawa et al. 2017). These subtle, yet critical differences may underlie sex differences in when and how aggression is deployed by changing how sensory information is integrated within these circuits.

There can also be sex differences in how neural circuits respond to the environment, rather than in the neural circuits that directly produce behavior. For example, analyses of neural circuits modulating social approach in California mice reveal few sex differences. In unstressed males and females, oxytocin receptors within the nucleus accumbens (Williams et al. 2020) and vasopressin V1a receptors in the bed nucleus of the stria terminalis (BNST) promote social approach. Indeed, across numerous analyses of gene expression or neuropeptides in the nucleus accumbens (Campi et al. 2014) and BNST (Greenberg et al. 2014; Duque-Wilckens et al. 2016, 2018; Steinman et al. 2016), few sex differences are observed. However social defeat induces sex-specific effects in these circuits (Steinman et al. 2019). For example, social defeat increases oxytocin synthesis as well as the reactivity of oxytocin neurons within the BNST (Steinman et al. 2016). Oxytocin synthesis in the BNST is necessary for stress-induced vigilance in female California mice, whereas oxytocin infusions into the BNST are sufficient to increase vigilance and reduce social approach in both males and females (Duque-Wilckens et al. 2020). Thus, it appears that a sex difference in the effects of stress on oxytocin neurons within the BNST drives sex differences in social approach and vigilance. Sex differences in transcriptional responses to social defeat have also been observed in the nucleus accumbens (Hodes et al. 2015; AV Williams et al., in prep.).

These sex-specific effects of stress on transcription within the brain contribute to important sex differences in behavioral responses following stress.

Beyond looking for "which" brain regions might have a role in mediating sex differences, comprehensive behavioral phenotyping will also allow us to assess "when" sex differences in behavior occurs. Sex differences in behavior may be seasonal (Henningsen et al. 2016) or may have different trajectories of expression across the life span (Schulz and Sisk 2016; Choleris et al. 2018). Collecting video data at different time points is now relatively low cost, and may reveal new insights about development.

CONCLUSION: SEX DIFFERENCES ARE "INDIVIDUAL" DIFFERENCES

The arc of neuroscience research is starting to bend back toward behavior. The development and distribution of new tools for animal pose tracking, in conjunction with the ability to record and integrate other multimodal streams of information, now allow us to capture a more holistic picture of animal behavior in the laboratory than has ever been possible. Broad access to these tools and the adoption of evolving strategies for analyzing these data may propel the next generation of neuroscientific researchers to fully characterize behavior, including sex differences in behavior, before launching a study. Not only will this approach yield novel behavioral insights, enabling the discovery of new behaviors and internal states, it will prevent potential experimental pitfalls, including oversimplistic single-variable readouts for complex states. Last, understanding the richness of behavioral variability (individual differences) is a long sought-after experimental goal. The quest to understand sex differences in behavior should be framed in this light, rather than in the rather narrow and often inaccurate confines of "maleness" and "femaleness." Striving to constantly and thoroughly evaluate population-wide individual variability in behavior will pay dividends in the ability to interpret neuroscientific data, and to design insightful behavioral readouts of complex internal and affective states.

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a039164

Cold Spring Harbor Perspectives in Biology

www.cshperspectives.org

REFERENCES

Anderson DJ, Perona P. 2014. Toward a science of computational ethology. *Neuron* **84:** 18–31. doi:10.1016/j.neuron.2014.09.005

Ball GF, Balthazart J, McCarthy MM. 2014. Is it useful to view the brain as a secondary sexual characteristic? *Neurosci Biobehav Rev* **46:** 628–638. doi:10.1016/j.neubiorev.2014.08.009

Becker JB, Prendergast BJ, Liang JW. 2016. Female rats are not more variable than male rats: a meta-analysis of neuroscience studies. *Biol Sex Differ* **7:** 34. doi:10.1186/s13293-016-0087-5

Berman GJ, Choi DM, Bialek W, Shaevitz JW. 2014. Mapping the stereotyped behaviour of freely moving fruit flies. *J R Soc Interface* **11:** 20140672. doi:10.1098/rsif.2014.0672

Bernstein IS. 1991. An empirical comparison of focal and ad libitum scoring with commentary on instantaneous scans, all occurrence and one-zero techniques. *Anim Behav* **42:** 721–728. doi:10.1016/S0003-3472(05)80118-6

Campi KL, Greenberg GD, Kapoor A, Ziegler TE, Trainor BC. 2014. Sex differences in effects of dopamine D1 receptors on social withdrawal. *Neuropharmacology* **77:** 208–216. doi:10.1016/j.neuropharm.2013.09.026

Choleris E, Galea LAM, Sohrabji F, Frick KM. 2018. Sex differences in the brain: implications for behavioral and biomedical research. *Neurosci Biobehav Rev* **85:** 126–145. doi:10.1016/j.neubiorev.2017.07.005

Clark JW, Drummond SPA, Hoyer D, Jacobson LH. 2019. Sex differences in mouse models of fear inhibition: fear extinction, safety learning, and fear-safety discrimination. *Br J Pharmacol* **176:** 4149–4158. doi:10.1111/bph.14600

Coffey KR, Marx RG, Neumaier JF. 2019. DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. *Neuropsychopharmacology* **44:** 859–868.

Darwin C. 1882. *The origin of species*. John Murray, London.

Datta SR, Anderson DJ, Branson K, Perona P, Leifer A. 2019. Computational neuroethology: a call to action. *Neuron* **104:** 11–24. doi:10.1016/j.neuron.2019.09.038

De Vries GJ, Boyle PA. 1998. Double duty for sex differences in the brain. *Behav Brain Res* **92:** 205–213.

Dolensek N, Gehrlach DA, Klein AS, Gogolla N. 2020. Facial expressions of emotion states and their neuronal correlates in mice. *Science* **368:** 89–94. doi:10.1126/science.aaz9468

Duque-Wilkens N, Steinman MQ, Laredo SA, Hao R, Perkeybile AM, Bales KL, Trainor BC. 2016. Inhibition of vasopressin V1a receptors in the medioventral bed nucleus of the stria terminalis has sex- and context-specific anxiogenic effects. *Neuropharmacology* **110:** 59–68. doi:10.1016/j.neuropharm.2016.07.018

Duque-Wilkens N, Steinman MQ, Busnelli M, Chini B, Yokoyama S, Pham M, Laredo SA, Hao R, Perkeybile AM, Minie VA, et al. 2018. Oxytocin receptors in the anteromedial bed nucleus of the stria terminalis promote stress-induced social avoidance in female California mice. *Biol Psychiatry* **83:** 203–213. doi:10.1016/j.biopsych.2017.08.024

Duque-Wilkens N, Torres LY, Yokoyama S, Minie VA, Tran AM, Petkova SP, Hao R, Ramos-Maciel S, Rios RA, Jackson K, et al. 2020. Extra-hypothalamic oxytocin neurons drive stress-induced social vigilance and avoidance. *Proc Natl Acad Sci* **117:** 26406–26413. doi:10.1073/pnas.2011890117

Emerson SB, Boyd SK. 1999. Mating vocalizations of female frogs: control and evolutionary mechanisms. *Brain Behav Evol* **53:** 187–197. doi:10.1159/000006594

Falkner AL, Dollar P, Perona P, Anderson DJ, Lin D. 2014. Decoding ventromedial hypothalamic neural activity during male mouse aggression. *J Neurosci* **34:** 5971–5984. doi:10.1523/JNEUROSCI.5109-13.2014

Falkner AL, Wei D, Song A, Watsek LW, Chen I, Chen P, Feng JE, Lin D. 2020. Hierarchical representations of aggression in a hypothalamic-midbrain circuit. *Neuron* **106:** 637–648.e6. doi:10.1016/j.neuron.2020.02.014

Forgie ML, Stewart J. 1993. Sex differences in amphetamine-induced locomotor activity in adult rats: Role of testosterone exposure in the neonatal period. *Pharmacol Biochem Behav* **46:** 637–645.

Friard O, Gamba M. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol Evol* **7:** 1325–1330. doi:10.1111/2041-210X.12584

Goffinet J, Brudner S, Mooney R, Pearson J. 2021. Inferring low-dimensional latent descriptions of animal vocalizations. *eLife* **10:** e67855. doi:10.7554/eLife.67855

Golden SA, Covington HE III, Berton O, Russo SJ. 2011. A standardized protocol for repeated social defeat stress in mice. *Nat Protoc* **6:** 1183–1191. doi:10.1038/nprot.2011.361

Goodwill HL, Manzano-Nieves G, Gallo M, Lee HI, Oyerinde E, Serre T, Bath KG. 2019. Early life stress leads to sex differences in development of depressive-like outcomes in a mouse model. *Neuropsychopharmacology* **44:** 711–720. doi:10.1038/s41386-018-0195-5

Graving JM, Chae D, Naik H, Li L, Koger B, Costelloe BR, Couzin ID. 2019. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife* **8:** e47994. doi:10.7554/eLife.47994

Greenberg GD, Laman-Maharg A, Campi KL, Voigt H, Orr VN, Schaal L, Trainor BC. 2014. Sex differences in stress-induced social withdrawal: role of brain derived neurotrophic factor in the bed nucleus of the stria terminalis. *Front Behav Neurosci* **7:** 223. doi:10.3389/fnbeh.2013.00223

Grill HJ, Norgren R. 1978. The taste reactivity test. I: Mimetic responses to gustatory stimuli in neurologically normal rats. *Brain Res* **143:** 263–279. doi:10.1016/0006-8993(78)90568-1

Gruene TM, Flick K, Stefano A, Shea SD, Shansky RM. 2015. Sexually divergent expression of active and passive conditioned fear responses in rats. *eLife* **4:** e11352. doi:10.7554/eLife.11352

Hammerschmidt K, Radyushkin K, Ehrenreich H, Fischer J. 2012. The structure and usage of female and male mouse ultrasonic vocalizations reveal only minor differences. *PLoS ONE* **7:** e41133. doi:10.1371/journal.pone.0041133

Hashikawa K, Hashikawa Y, Tremblay R, Zhang J, Feng JE, Sabol A, Piper WT, Lee H, Rudy B, Lin D. 2017. Esr1[+] cells in the ventromedial hypothalamus control female aggression. *Nat Neurosci* **20:** 1580–1590. doi:10.1038/nn.4644

B.C. Trainor and A.L. Falkner

Henningsen JB, Poirel VJ, Mikkelsen JD, Tsutsui K, Simonneaux V, Gauer F. 2016. Sex differences in the photoperiodic regulation of RF-amide related peptide (RFRP) and its receptor GPR147 in the Syrian hamster. *J Comp Neurol* **524:** 1825–1838. doi:10.1002/cne.23924

Hodes GE, Pfau ML, Purushothaman I, Ahn HF, Golden SA, Christoffel DJ, Magida J, Brancato A, Takaashi A, Flanigan ME, et al. 2015. Sex differences in nucleus accumbens transcriptome profiles associated with susceptibility versus resilience to subchronic variable stress. *J Neurosci* **35:** 16362–16376. doi:10.1523/JNEUROSCI.1392-15.2015

Hoffman EJ, Turner KJ, Fernandez JM, Cifuentes D, Ghosh M, Ijaz S, Jain RA, Kubo F, Bill BR, Baier H, et al. 2016. Estrogens suppress a behavioral phenotype in zebrafish mutants of the autism risk gene, *CNTNAP2. Neuron* **89:** 725–733. doi:10.1016/j.neuron.2015.12.039

Holy TE, Guo Z. 2005. Ultrasonic songs of male mice. *PLoS Biol* **3:** e386. doi:10.1371/journal.pbio.0030386

Hsu AI, Yttri EA. 2021. An open source unsupervised algorithm for identification and fast prediction of behaviors. https://www.biorxiv.org/content/10.1101/770271v3 [accessed August 27, 2021].

Hubbard DT, Blanchard DC, Yang M, Markham CM, Gervacio A, Chun-I L, Blanchard RJ. 2004. Development of defensive behavior and conditioning to cat odor in the rat. *Physiol Behav* **80:** 525–530. doi:10.1016/j.physbeh.2003.10.006

Kelly KK. 1993. Androgenic induction of brain sexual dimorphism depends on photoperiod in meadow voles. *Physiol Behav* **53:** 245–249.

Khan MH, McDonagh J, Khan S, Shahabuddin M, Arora A, Khan FS, Shao L, Tzimiropoulos G. 2020. AnimalWeb: a large-scale hierarchical dataset of annotated animal faces. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6939–6948.

Klibaite U, Berman GJ, Cande J, Stern DL, Shaevitz JW. 2017. An unsupervised method for quantifying the behavior of paired animals. *Physical Biol* **14:** 015006. doi:10.1088/1478-3975/aa5c50

Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D. 2017. Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93:** 480–490. doi:10.1016/j.neuron.2016.12.041

Laman-Maharg A, Trainor BC. 2017. Stress, sex and motivated behaviors. *J Neurosci Res* **95:** 83–92. doi:10.1002/jnr.23815

Li S, Günel S, Ostrek M, Ramdya P, Fua P, Rhodin H. 2020. Deformation-aware unpaired image translation for pose estimation on laboratory animals. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 13158–13168.

Lin D, Boyle MP, Dollar P, Lee H, Lein ES, Perona P, Anderson DJ. 2011. Functional identification of an aggression locus in the mouse hypothalamus. *Nature* **470:** 221–226. doi:10.1038/nature09736

Lischinsky JE, Lin D. 2020. Neural mechanisms of aggression across species. *Nat Neurosci* **23:** 1317–1328. doi:10.1038/s41593-020-00715-2

Mamlouk GM, Dorris DM, Barrett LR, Meitzen J. 2020. Sex bias and omission in neuroscience research is influenced by research model and journal, but not reported NIH funding. *Front Neuroendocrinol* **57:** 100835.

Maren S, De Oca B, Fanselow MS. 1994. Sex differences in hippocampal long-term potentiation (LTP) and Pavlovian fear conditioning in rats: positive correlation between LTP and contextual learning. *Brain Res* **661:** 25–34. doi:10.1016/0006-8993(94)91176-2

Martin P, Bateson P. 1994. *Measuring behavior, an introductory guide*, 2nd ed. Cambridge University Press, Cambridge.

Marr D, Nishihara HK. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc R Soc Lond B Biol Sci* **200:** 269–294.

Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Weygandt Mathis M, Bethge M. 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* **21:** 1281–1289. doi:10.1038/s41593-018-0209-y

Mathis A, Schneider S, Lauer J, Weygandt MM. 2020. A primer on motion capture with deep learning: principles, pitfalls, and perspectives. *Neuron* **108:** 44–65. doi:10.1016/j.neuron.2020.09.017

McCann KE, Huhman KL. 2012. The effect of escapable versus inescapable social defeat on conditioned defeat and social recognition in Syrian hamsters. *Physiol Behav* **105:** 493–497. doi:10.1016/j.physbeh.2011.09.009

Neunuebel JP, Taylor AL, Arthur BJ, Roian Egnor SE. 2015. Female mice ultrasonically interact with males during courtship displays. *eLife* **4:** e06203. doi:10.7554/eLife.06203

Newman EL, Covington HE, Suh J, Bicakci MB, Ressler KJ, DeBold JF, Miczek KA. 2019. Fighting females: neural and behavioral consequences of social defeat stress in female mice. *Biol Psychiatry* **86:** 657–668.

Nilsson SRO, Goodwin NL, Choong JJ, Hwang S, Wright HR, Norville ZC, Tong X, Lin D, Bentzley BS, Eshel N, et al. 2020. Simple behavioral analysis (SimBA)—an open source toolkit for computer classification of complex social behaviors in experimental animals. bioRxiv doi:10.1101/2020.04.19.049452

Odom KJ, Hall ML, Riebel K, Omland KE, Langmore NE. 2014. Female song is widespread and ancestral in songbirds. *Nat Commun* **5:** 3379. doi:10.1038/ncomms4379

Penna M, Capranica RR, Somers J. 1992. Hormone-induced vocal behavior and midbrain auditory sensitivity in the Green treefrog, *Hyla cinerea. J Comp Physiol A* **170:** 73–82. doi:10.1007/BF00190402

Pereira TD, Aldarondo DE, Willmore L, Kislin M, Wang SSH, Murthy M, Shaevitz JW. 2019. Fast animal pose estimation using deep neural networks. *Nat Methods* **16:** 117–125. doi:10.1038/s41592-018-0234-5

Pereira TD, Shaevitz JW, Murthy M. 2020. Quantifying behavior to understand the brain. *Nat Neurosci* **23:** 1537–1549. doi:10.1038/s41593-020-00734-z

Pernold K, Iannello F, Low BE, Rigamonti M, Rosati G, Scavizzi F, Wang J, Raspa M, Wiles MV, Ulfhake B. 2019. Towards large scale automated cage monitoring—diurnal rhythm and impact of interventions on in-cage activity of C57BL/6J mice recorded 24/7 with a non-disrupting capacitive-based technique. *PLoS ONE* **14:** e0211063. doi:10.1371/journal.pone.0211063

Perrot-Sinal TS, Kostenuik MA, Ossenkopp K-P, Kavaliers M. 1996. Sex differences in performance in the Morris

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a039164

water maze and the effects of initial nonstationary hidden platform training. *Behav Neurosci* **110:** 1309–1320.

Prendergast BJ, Onishi KG, Zucker I. 2014. Female mice liberated for inclusion in neuroscience and biomedical research. *Neurosci Biobehav Rev* **40:** 1–5. doi:10.1016/j.neubiorev.2014.01.001

Pryce CR, Lehmann J, Feldon J. 1999. Effect of sex on fear conditioning is similar for context and discrete CS in Wistar, Lewis and Fischer rat strains. *Pharmacol Biochem Behav* **64:** 753–759. doi:10.1016/S0091-3057(99)00147-1

Roth ME, Carroll ME. 2004. Sex differences in the acquisition of IV methamphetamine self-administration and subsequent maintenance under a progressive ratio schedule in rats. *Psychopharmacology (Berl)* **172:** 443–449. doi:10.1007/s00213-003-1670-0

Schulz KM, Sisk CL. 2016. The organizing actions of adolescent gonadal steroid hormones on brain and behavioral development. *Neurosci Biobehav Rev* **70:** 148–158. doi:10.1016/j.neubiorev.2016.07.036

Segalin C, Williams J, Karigo T, Hui M, Zelikowsky M, Sun JJ, Perona P, Anderson DJ, Kennedy A. 2020. The mouse action recognition system (MARS): a software pipeline for automated analysis of social behaviors in mice. bioRxiv doi:10.1101/2020.07.26.222299

Steinman MQ, Duque-Wilckens N, Greenberg GD, Hao R, Campi KL, Laredo SA, Laman-Maharg A, Manning CE, Doig IE, Lopez EM, et al. 2016. Sex-specific effects of stress on oxytocin neurons correspond with responses to intranasal oxytocin. *Biol Psychiatry* **80:** 406–414. doi:10.1016/j.biopsych.2015.10.007

Steinman MQ, Duque-Wilckens N, Trainor BC. 2019. Complementary neural circuits for divergent effects of oxytocin: social approach versus social anxiety. *Biol Psychiatry* **85:** 792–801. doi:10.1016/j.biopsych.2018.10.008

Sturman O, von Ziegler L, Schläppi C, Akyol F, Grewe B, Privitera M, Slominski D, Grimm C, Thieren L, Zerbi V, et al. 2020. Deep learning-based behavioral analysis reaches human accuracy and is capable of outperforming commercial solutions. *Neuropsychopharmacology* **45:** 1942–1952. doi:10.1038/s41386-020-0776-y

Tobias ML, Marin ML, Kelley DB. 1993. The roles of sex, innervation, and androgen in laryngeal muscle of *Xenopus laevis*. *J Neurosci* **13:** 324–333. doi:10.1523/JNEUROSCI.13-01-00324.1993

Trainor BC, Pride MC, Villalon Landeros R, Knoblauch NW, Takahashi EY, Silva AL, Crean KK. 2011. Sex differences in social interaction behavior following social defeat stress in the monogamous California mouse (*Peromyscus californicus*). *PLoS ONE* **6:** e17405. doi:10.1371/journal.pone.0017405

Trainor BC, Takahashi EY, Campi KL, Florez SA, Greenberg GD, Laman-Maharg A, Laredo SA, Orr VN, Silva AL,

Steinman MQ. 2013. Sex differences in stress-induced social withdrawal: independence from adult gonadal hormones and inhibition of female phenotype by corncob bedding. *Horm Behav* **63:** 543–550. doi:10.1016/j.yhbeh.2013.01.011

Van Segbroeck M, Knoll AT, Levitt P, Narayanan S. 2017. MUPET-mouse ultrasonic profile ExTraction: a signal processing tool for rapid and unsupervised analysis of ultrasonic vocalizations. *Neuron* **94:** 465–485.e5. doi:10.1016/j.neuron.2017.04.005

Van Swearingen AED, Walker QD, Kuhn CM. 2013. Sex differences in novelty- and psychostimulant-induced behaviors of C57BL/6 mice. *Psychopharmacology (Berl)* **225:** 707–718. doi:10.1007/s00213-012-2860-4

Vogel AP, Tsanas A, Scattoni ML. 2019. Quantifying ultrasonic mouse vocalizations using acoustic analysis in a supervised statistical machine learning framework. *Sci Rep* **9:** 8100. doi:10.1038/s41598-019-44221-3

von Ziegler L, Sturman O, Bohacek J. 2021. Big behavior: challenges and opportunities in a new era of deep behavior profiling. *Neuropsychopharmacology* **46:** 33–44.

Warren BL, Mazei-Robison MS, Robison AJ, Iñiguez SD. 2020. Can I get a witness? Using vicarious defeat stress to study mood-related illnesses in traditionally understudied populations. *Biol Psychiatry* **88:** 381–391. doi:10.1016/j.biopsych.2020.02.004

Wilczynski W, Quispe M, Muñoz MI, Penna M. 2017. Arginine vasotocin, the social neuropeptide of amphibians and reptiles. *Front Endocrinol* **8:** 186. doi:10.3389/fendo.2017.00186

Williams AV, Duque-Wilckens N, Ramos-Maciel S, Campi KL, Bhela SK, Xu CK, Jackson K, Chini B, Pesavento PA, Trainor BC. 2020. Social approach and social vigilance are differentially regulated by oxytocin receptors in the nucleus accumbens. *Neuropsychopharmacology* **45:** 1423–1430. doi:10.1038/s41386-020-0657-4

Wiltschko AB, Tsukahara T, Zeine A, Anyoha R, Gillis WF, Markowitz JE, Peterson RE, Katon J, Johnson MJ, Datta SR. 2020. Revealing the structure of pharmacobehavioral space through motion sequencing. *Nat Neurosci* **23:** 1433–1443.

Wright EC, Hostinar CE, Trainor BC. 2020. Anxious to see you: neuroendocrine mechanisms of social vigilance and anxiety during adolescence. *Eur J Neurosci* **52:** 2516–2529. doi:10.1111/ejn.14628

Yohn CN, Dieterich A, Bazer AS, Maita I, Giedraitis M, Samuels BA. 2019. Chronic non-discriminatory social defeat is an effective chronic stress paradigm for both male and female mice. *Neuropsychopharmacology* **44:** 2220–2229. doi:10.1038/s41386-019-0520-7

Cold Spring Harbor
Perspectives in Biology

Quantifying Sex Differences in Behavior in the Era of "Big" Data

Brian C. Trainor and Annegret L. Falkner

Cold Spring Harb Perspect Biol published online October 4, 2021

Subject Collection   Sex Differences in Brain and Behavior