Inclusion of females does not increase variability in rodent research studies

Annaliese K Beery[1,2]

Underrepresentation of female subjects in animal research has gained attention in recent years, and new NIH guidelines aim to address this problem early, at the grant proposal stage. Many researchers believe that requirements regarding use of females will hamper research because of a need for increased sample sizes, and increased costs. Empirical research across multiple rodent species and traits demonstrates that females are not more variable than males, and that for most traits, female estrous cyclicity need not be considered. Statistical simulations, presented here, illustrate how factorial designs can reduce the need for additional research subjects, and cultural issues around the inclusion of male and female subjects in research are discussed.

Addresses
[1] Department of Psychology, Department of Biology, Neuroscience Program, Smith College, Northampton, MA, USA
[2] Neuroscience and Behavior Graduate Program, University of Massachusetts, Amherst, MA, USA

Corresponding author: Beery, Annaliese K. (abeery@smith.edu)

Introduction

Studies incorporating females have revealed marked differences in basic biological processes such as pain signaling [1,2••]. Female subjects are underrepresented in animal research across disciplines, however [3••], and lack of pre-clinical research on female subjects has likely resulted in poorer treatment outcomes for women [4,5]. In 2014, noting potential human health consequences of this research bias, the NIH instituted policies to encourage use of both male and female animal research subjects, and consideration of sex as a biological variable [6••,7]. Biological sex – classification as generally male or female based on genetic and physiological features – is typically distinguished from gender – one's self-representation as male, female, or non-binary. Inclusion of both sexes in animal research studies should drive important discoveries in both basic and clinically relevant research [5].

The call for inclusion of females has met with some criticism [8,9]. One oft-used justification for focusing on males is that females are presumed to be more variable, in part due to estrous cyclicity. Other pushback comes from concern that where sex differences exist, use of males and females will reduce statistical power because of greater spread of pooled data or smaller sub-samples of each sex [8]. Finally, some are concerned that increased attention paid to sex differences in preclinical studies will overemphasize what are sometimes small differences in the midst of fundamental similarities, or fail to model sex/gender differences in human health that may have important sociocultural components [9].

This review describes the extent of sex bias and presents the findings of empirical studies and analyses that address these concerns surrounding the inclusion of female subjects.

## How bad is the *status quo*?
Sex bias
The use of predominantly male animal research subjects has been documented in many fields. In a survey across biological disciplines, we found male bias in 8 of 10 fields (general biology, neuroscience, physiology, pharmacology, endocrinology, behavioral physiology, behavior, and zoology — reproduction and immunology were the exceptions) [3••]. Similar bias toward the use of male subjects has been found in surveys of preclinical animal research on pain, cardiovascular disease, diabetes, and surgical methods [10–13]. In the surgical literature, 80% of studies that specified sex used only male subjects [11].

We found neuroscience to be one of the worst offenders, with over five studies on males for each study on females; only ~20% of studies used both sexes, and 25% did not specify the sex of study subjects (Figure 1a). Lack of reporting of subject sex has been documented at similar rates (20, 22, and 26%) in other surveys [11,13,14]. Even when subject sex is reported, it is sometimes not evident until late in the results, or may require accessing online supplementary material. Subsequent analysis of the neuroscience literature suggests that omission of subject sex has decreased in recent years, but the number of male-only rodent studies has increased, and analysis by subject sex in mixed-sex studies remains infrequent [15].

The bias in animal subject use does not reflect differences in rates of men and women presenting with

Figure 1
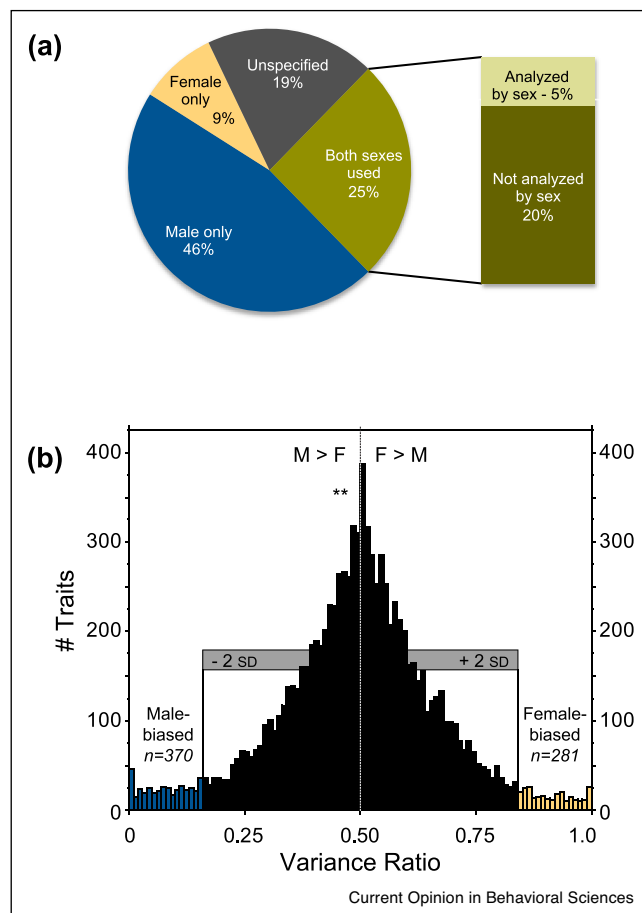
Data on sex bias and trait variability. **(a)** Inclusion of male and female animal research subjects was surveyed across 10 biological disciplines (neuroscience shown here). Even when both sexes were used, analyses typically did not consider subject sex. Analysis of data from Ref. [3••]. **(b)** Coefficients of variation were assessed in male and female mice across >9900 measurements of traits. Variability was similar in males and females, with more male-biased than female-biased traits, and a mean variance ratio significantly lower than 0.5. Modified (with permission) from Ref. [22••].

diseases or disorders; the percentage of women diagnosed with a given condition exceeded the percentage of non-human female subjects in research studies on that condition, in each of several disorders sampled [16]. However, the gender of study authors appears to play a role, as female authorship was significantly positively correlated with the inclusion of both sexes and analysis by subject gender/sex [17].

Analysis by subject sex is not the norm

Even when studies include males and females, analysis by subject sex is infrequent. In a survey of surgical research on animals, only 1% of papers analyzed results by sex [11]. In the aforementioned neuroscience example, 5.5% of the human and animal research papers sampled used both sexes and analyzed results by sex (Figure 1a); other fields showed even lower rates [3••].

In 1993, the NIH introduced the Revitalization Act, requiring that women be included in NIH-funded clinical research [18], and requiring sub-group analysis by sex to be enabled. A 2011 analysis found that less than a third of studies that were required by the NIH to analyze results by subject sex/gender had published analysis by this factor [19••]. Similar absences of mandatory subgroup analyses have occurred in drug safety reporting to the FDA [20••]. In November 2017, the NIH announced an amendment [NOT-OD-18-014] to the NIH Policy and Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research, stipulating that results of valid subgroup analysis (including by subject sex/gender) must be submitted to clinical-trials.gov. At present, the NIH instructs preclinical investigators to report subject sex and consider sex as a biological variable, but analysis of sex differences need not be performed [7]. Surveys of researchers revealed that how subject sex should best be considered in analysis is not uniformly agreed upon, and that researcher discretion in selecting analyses appropriate to the sample is preferred [21].

Why not use both sexes? Countering assumptions surrounding use of females

Although good reasons exist to study females or males alone, the rationale for use of only one sex often stems from unfounded concerns about the use of females or both sexes. These assumptions: that females are more variable than males, that females must be tested across the estrous cycle, and that inclusion of both sexes increases variability, are each countered below.

In considering variability, some common statistical terms and principles will be of service. Variation in a data set can be described in terms of *standard deviation* (SD), a measure of the dispersion or spread of data. The *coefficient of variation* describes this standard deviation as a fraction of the mean value (SD/mean) so that it is scaled relative to the data. For example, a sample with mean $\pm$ SD of 1 ± 0.05 and one with 1000 ± 50 have equivalent levels of variability relative to their respective means. *Statistical power* refers to the probability of correctly rejecting the null hypothesis, for instance to detect differences between samples at specific conventionally defined statistical thresholds (i.e. $p < 0.05$). For a given sample size, power is greater when mean differences between groups are higher. For a given mean difference between groups, power is greater when sample sizes are higher, allowing better estimation of differences in sample means.

Assumption 1: females are more variable than males

Even if females were more variable than males, it would not alter the *importance* of studying both sexes, only the difficulty. Fortunately, researchers do not have to contend with that scenario, as recent empirical analyses of variability in males and females show that unstaged females are not more variable than males across diverse traits, from gene expression to hormone levels in multiple species [22••,23••,24•,25–27]. Indeed, male mice appeared to exhibit slightly, but significantly, greater mean variability than female mice (Figure 1b), with substantially higher coefficients of variation (CVs) for hormone measures, metabolism-related traits, and morphology. Females did not have significantly greater variability in any category [22••]. Slightly greater male variability was also found in analysis of microarray datasets in both mice and humans [25]. In rats no sex difference in overall variability was found, but males exhibited significantly higher variability on nervous system measures including neurochemistry and electrophysiology measures, while females exhibited higher variability in the category of non-brain measures [23••]. In hamsters, no significant differences in variability were found between male, female, and ovariectomized (non-cycling) females [27].

Assumption 2: females must be tested across the estrous cycle

Predictable variation in hormone circulation across the estrous cycle contributes to some variation in physiology and behavior, but how much? The similar variability between males and *unstaged* females, described above, suggests estrous variability is no greater than intrinsic variability in males. But perhaps similar overall variability comes from different sources: it is plausible that females might exhibit consistent estrous-cycle dependent variability in several traits, while males exhibit variability on different time-scales, or more variability between individuals. In-depth analysis of a particular trait (body temperature in mice) revealed equivalent overall variability in males and females, but different time-scales of variability [24•]. Analysis of a variety of traits tested over 2 consecutive estrous cycles in intact female, ovariectomized female, and male hamsters, revealed no consistent timing differences across traits. Furthermore, staging results by phase of the estrous cycle offered no reduction in variability, even for traits such as sucrose preference where estrous cycle variation can be detected under circumstances that are optimized to find it [27]. Thus, testing across the estrous cycle can be considered a specific tool for use in a small set of specific instances rather than a necessary procedure in most studies [28].

While estrous cyclicity is not a major source of variability, other documented factors may provide avenues for researchers to reduce variability and increase statistical power, including the number of animals/cage and rodent strain [22••,23••]. Consideration of frequently overlooked study details including bedding type [29], biological sex of the experimenter [30], and overlap of animal shipment (a stressor) with puberty [31,32], may also increase consistency across studies. The existence of substantial differences in findings between laboratories despite careful efforts at duplication of conditions suggests that more factors may be important than currently realized [33].

Assumption 3: use of both sexes reduces statistical power and slows progress

Thus far we have considered variability of males and females relative to each group. For traits in which a robust sex difference exists, however, pooling males and females in one group would increase the variability around a combined mean. This gives rise to the concern that sample sizes might need to be doubled to identify treatment effects in studies using both females and males [8]. Factorial designs, however, can evaluate main effects of treatment and subject sex with effectively the same power as pair-wise tests, without increased sample size [34]. Additional factors can also be added at the same sample size with approximately the same power, provided the effect sizes of each new factor are no smaller than the effect sizes used to generate the original estimate of the number of subjects needed [34,35].

Illustrations of the statistical ramifications of analyzing results with subject sex as a factor are explored in greater detail in Figure 2 and Box 1. The only circumstance in which a notable reduction in power occurs is when there is an interaction between treatment and sex — which is to say that males and females respond very differently to the treatment. In that case, follow-up testing that is methodologically designed to capture sex differences and their origins [28] will be biologically meaningful and important. The factorial approach is powerful, but not without potential weaknesses. ANOVA on sex*treatment generates three F-values without compensation for three tests, leading to higher type I error rate than one *t*-test (explaining why the 'treatment' factor performs as well as the *t*-test in scenario 1). This is important to keep in mind if additional factors are added. Also, while scenario 3 is extreme, more intermediate interactions will be less easily detected. If assessing sex differences is a primary rather than secondary goal, increased sample size will improve detection of interactions. Larger samples always provide better potential for analysis, and many have called for fewer studies performed on more individuals, particularly in neuroscience [36]. Sub-division of this sample into males and females and use of factorial designs is an effective method of analysis both in theory and in practice [37].

Much ado about sex

Sex differences can be small or large, insignificant or critical. One concern about the reporting of sex


Box 1 Including females does not necessitate larger sample sizes

Some researchers are concerned that use of males and females will result in increased variance, reduced power, and the need to test more animals [8]. Fortunately, factorial analysis methods result in little to no loss of statistical power [34,35], except when there is a sizable interaction between sex and treatment. In that case, it is especially important to study both sexes initially, and subsequent, sex-specific analysis or studies may be needed.

For illustration, simulated outcomes of 2-group and factorial tests are presented. Consider an experiment with treatments 'A' and 'B' (e.g. hot vs. cold room temperature) and an outcome measure (e.g. distance traveled). Results are shown for tests using samples of 12 females (f), 12 males (m), or 6f and 6m — in the presence and absence of sex differences and interaction effects. In each scenario there is a ±5% change in mean and SD in each sex.

In scenario 1 there is no sex difference (Figure 2a). This is common, and results in no cost to mixing sexes. All analyses yield equivalent effects of treatment, and 2-way ANOVA on sex*treatment indicates no interaction. In scenario 2 there is a large sex difference and a moderate treatment difference (Figure 2b). This is an oft-feared scenario in which pooling males and females reduces statistical power in t-tests across treatments. 2-way ANOVA results in no loss of power to detect treatment effects, however, as the test detects treatment effects within each sub-group. Finally, scenario 3 (Figure 2c) represents a possible 'worst-case scenario' with a large treatment effect in females and an equally sizable *but opposite* effect in males. This results in the eradication of a treatment difference, but the presence of a strong ANOVA interaction effect signals that sex-specific follow-up is strongly indicated. R code for these simulations is available at https://osf.io/6q73b/.

Figure 2

Current Opinion in Behavioral Sciences
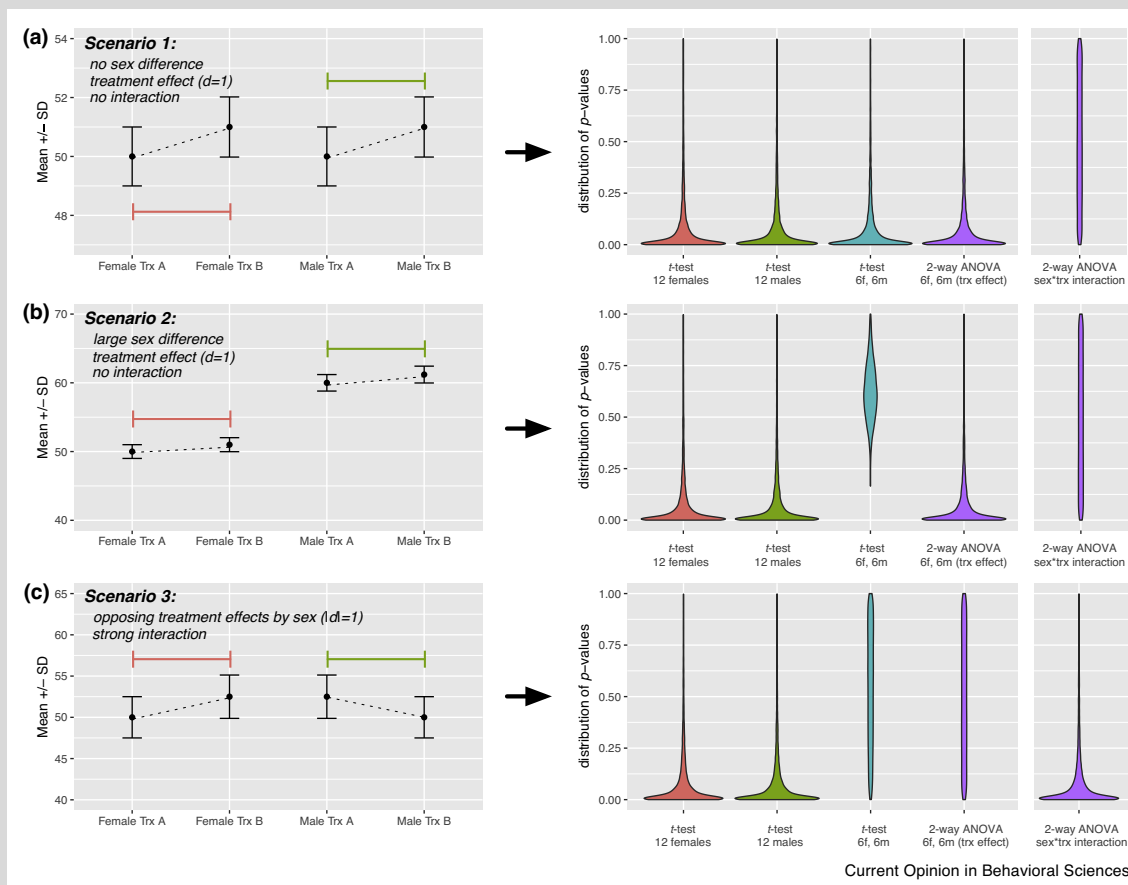
p-value distributions for simulated group compositions and treatment effects. **Left panels**: mean ± SD of each group, used to generate 10 000 normally distributed samples of subjects (12f, 12m, or 6f and 6m) in each treatment. Coefficients of variation (SD/mean) were matched for each sex/treatment combination, as were effect sizes for treatment comparisons in all-male or female groups (Cohen's d = 1 using SD of lower group; .97–.99 using pooled SD). **Right panels**: violin plots of p-values from t-tests between single- and mixed-sex groups in treatments A versus B, or from the treatment factor from 2-way sex*treatment ANOVA on mixed-sex groups. The fifth distribution consists of p-values from the ANOVA's interaction term across runs. Even with a large sex difference, there is no loss from testing half males and half females when a factorial analysis is used, as long as there is no interaction. When a strong interaction is present, factorial analysis cannot detect a unified treatment effect, but the interaction effect indicates that subgroup analysis by sex and possible follow-up experiments are merited.

---

differences in animal studies is that it may lead to overestimation of human sex/gender differences, especially in brain and behavior [9]. One action researchers can take to address this concern is to indicate the effect size of any sex difference that is reported [5]. It is also critical that in translating findings to humans, we consider that rodents share only some traits with people, and that both sex and gender play a role in differences between women and men.

On the other hand, we should not underestimate the potential importance of sex differences. In some cases, small phenotypic differences stem from fundamental differences in underlying mechanisms. For example, male and female mice exhibit significant differences in pain sensitivity but have largely overlapping distributions. These 'small but significant' differences may arise in part based on pain modulatory pathways that differ with subject sex and hormone exposure [1,2••]. Sex-specific mechanisms of synaptic inhibition have been discovered in the hippocampus, demonstrating that even basic mechanisms of neuromodulation can vary with sex [38]. Multiple sex differences in the pathways underlying social behaviors in males and females have been documented, even when males and females exhibit similar behaviors, such as pair-bonding with a mate. One consequence of pathway differences underlying similar behaviors is that the same perturbations of neurochemicals or environmental factors can have opposing effects in males and females [39–41]. Recent work on parental behavior reveals that some pathways contribute similarly to both maternal and paternal behaviors in mice [42,43], while other circuits and neuropeptide expression patterns differ in important ways [44]. Perhaps a more important view is that studies assessing mechanisms underlying behavior in both sexes are so rare that we often have little idea of the magnitude and relevance of differences [45]. There are likely many more sex-specific differences in fundamental physiological processes we have missed.

Another concern is that preclinical research on sex differences may not be beneficial enough to humans to merit study, as some differences between men and women may be based on factors related more to culture and gender than to biological sex [9]. The concern here is that if there is a statistical or conceptual cost to including females, there needs to be a clear benefit [9]. While others have discussed benefits at length [5], this article seeks to mitigate the issue of cost. And many basic sex differences have already been discovered in humans. For instance, rodent research on sex differences in dopamine fiber densities and modulation of dopamine circuitry by estradiol [46] inspired research that discovered effects of endogenous fluctuation in estradiol in women on prefrontal cortex-dependent cognitive tasks [47]. The role of melanocortin receptor 1 in pain processing in female but not male mice was also recapitulated in humans, in whom only females with melanocortin 1 receptor mutations exhibit altered pain processing [1].

The discovery of differences such as these, despite limitations in the ability to assess mechanistic variation in humans *in vivo*, suggests that many more human sex differences in underlying mechanisms will ultimately be detected. While there is no guarantee that specific animal research findings or specific sex differences will translate to humans, this can be improved by conducting comparative research across species, as well as corresponding research in humans. The unacceptable alternative is that we adopt the rodent equivalent of considering 'man' to stand for 'human.'

Conflict of interest statement
Nothing declared.

Acknowledgements

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as

- • of special interest
- •• of outstanding interest

1. Mogil JS, Wilson SG, Chesler EJ, Rankin AL, Nemmani KVS, Lariviere WR, Groce MK, Wallace MR, Kaplan L, Staud R *et al.*: **The melanocortin-1 receptor gene mediates female-specific mechanisms of analgesia in mice and humans**. *Proc Natl Acad Sci U S A* 2003, **100**:4867-4872.

2. Sorge RE, Mapplebeck JCS, Rosen S, Beggs S, Taves S,
•• Alexander JK, Martin LJ, Austin J-S, Sotocinal SG, Chen D *et al.*: **Different immune cells mediate mechanical pain hypersensitivity in male and female mice**. *Nat Neurosci* 2015, **18**:1081-1083.
Prime example of a fundamental sex difference in a basic biological process discovered because this research group studies both males and females.

3. Beery AK, Zucker I: **Sex bias in neuroscience and biomedical**
•• **research**. *Neurosci Biobehav Rev* 2011, **35**:565-572.
Literature survey of the extent of sex bias across biological research domains in 2009 and historically. Species use over time is also documented in the online supplement.

4. Heinrich J: **Most drugs withdrawn in recent years had greater health risks for women**. *Gen Account Off-01-286R* 2001.

5. Klein SL, Schiebinger L, Stefanick ML, Cahill L, Danska J, de Vries GJ, Kibbe MR, McCarthy MM, Mogil JS, Woodruff TK *et al.*: **Opinion: sex inclusion in basic research drives discovery**. *Proc Natl Acad Sci* 2015.

6. Clayton JA, Collins FS: **Policy: NIH to balance sex in cell and**
•• **animal studies**. *Nature* 2014, **509**:282-283.
Announcement of new NIH guidelines regarding consideration of sex as a biological variable.

7. Clayton JA: **Studying both sexes: a guiding principle for biomedicine**. *FASEB J* 2016, **30**:519-524.

8. Fields RD: **NIH policy: mandate goes too far**. *Nature* 2014, **510**:340.

9. Richardson SS, Reiches M, Shattuck-Heidorn H, LaBonte ML, Consoli T: **Opinion: focus on preclinical sex differences will not address women's and men's health disparities**. *Proc Natl Acad Sci U S A* 2015, **112**:13419-13420.

10. Mogil JS, Chanda ML: **The case for the inclusion of female subjects in basic science studies of pain**. *Pain* 2005, **117**:1-5.

11. Yoon DY, Mansukhani NA, Stubbs VC, Helenowski IB, Woodruff TK, Kibbe MR: **Sex bias exists in basic science and translational surgical research**. *Surgery* 2014, **156**:508-516.

12. Ramirez FD, Motazedian P, Jung RG, Di Santo P, MacDonald Z, Simard T, Clancy AA, Russo JJ, Welch V, Wells GA: **Sex bias is increasingly prevalent in preclinical cardiovascular research: implications for translational medicine and health equity for women: a systematic assessment of leading cardiovascular journals over a 10-year period**. *Circulation* 2017, **135**:625-626.

13. Flórez-Vargas O, Brass A, Karystianis G, Bramhall M, Stevens R, Cruickshank S, Nenadic G: **Bias in the reporting of sex and age in biomedical research on mouse models**. *eLife* 2016, **5**.

14. Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, Fry D, Hutton J, Altman DG: **Survey of the quality of experimental design, statistical analysis and reporting of research using animals**. *PLoS ONE* 2009, **4**:e7824.

15. Will TR, Proaño SB, Thomas AM, Kunz LM, Thompson KC, Ginnari LA, Jones CH, Lucas S-C, Reavis EM, Dorris DM *et al.*: **Problems and progress regarding sex bias and omission in neuroscience research**. *eNeuro* 2017 http://dx.doi.org/10.1523/ENEURO.0278-17.2017.

16. Zucker I, Beery AK: **Males still dominate animal studies**. *Nature* 2010, **465**:690.

17. Nielsen MW, Andersen JP, Schiebinger L, Schneider JW: **One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis**. *Nat Hum Behav* 2017, **1**:791.

18. Congress US: *National Institutes of Health revitalization act of 1993*. Public Law; 1993.

19. Foulkes MA: **After inclusion, information and inference:**
•• **reporting on clinical trials results after 15 years of monitoring inclusion of women**. *J Womens Health 2002* **(20)**:2011:829-836.
Analysis of adherence to NIH guidelines regarding inclusion of females and analysis by gender/sex in NIH-funded phase 3 clinical trials.

20. Mazure CM, Jones DP: **Twenty years and still counting:**
•• **including women as participants and studying sex and gender in biomedical research**. *BMC Womens Health* 2015, **15**.
Comprehensive review of historical developments regarding sex and gender in research over the past 20 years, with particular attention paid to recent progress and continuing shortfalls.

21. Tannenbaum C, Schwarz JM, Clayton JA, de Vries GJ, Sullivan C: **Evaluating sex as a biological variable in preclinical research: the devil in the details**. *Biol Sex Differ* 2016, **7**:13.

22. Prendergast BJ, Onishi KG, Zucker I: **Female mice liberated for**
•• **inclusion in neuroscience and biomedical research**. *Neurosci Biobehav Rev* 2014, **40**:1-5.
Landmark literature survey of variability in mouse traits revealing that females are not more variable than males, even without tracking of estrous cycles. Other variables that influence trait variability are documented.

23. Becker JB, Prendergast BJ, Liang JW: **Female rats are not more**
•• **variable than male rats: a meta-analysis of neuroscience studies**. *Biol Sex Differ* 2016, **7**.
Similar analysis and findings as Prendergast*et al*. above, but in rats. Also documents housing and strain differences in variability.

24. Smarr BL, Grant AD, Zucker I, Prendergast BJ, Kriegsfeld LJ: **Sex**
• **differences in variability across timescales in BALB/c mice**. *Biol Sex Differ* 2017, **8**:7.
Analysis of sources of body temperature variability in males and females.

25. Itoh Y, Arnold AP: **Are females more variable than males in gene expression? Meta-analysis of microarray datasets**. *Biol Sex Differ* 2015, **6**.

26. Dayton A, Exner EC, Bukowy JD, Stodola TJ, Kurth T, Skelton M, Greene AS, Cowley AW: **Breaking the cycle: estrous variation does not require increased sample size in the study of female rats**. *Hypertens Dallas Tex 1979* 2016, **68**:1139-1144.

27. Moshofsky KJ, Beery AK: **The role of sex and ovarian cycles in rodent behavioral and physiological variability**. *Poster P122 21st Meet Soc Behav Neuroendocrinol* 2017 http://dx.doi.org/10.6084/m9.figshare.6753320.

28. Becker JB, Arnold AP, Berkley KJ, Blaustein JD, Eckel LA, Hampson E, Herman JP, Marts S, Sadee W, Steiner M *et al.*: **Strategies and methods for research on sex differences in brain and behavior**. *Endocrinology* 2005, **146**:1650-1673.

29. Landeros RV, Morisseau C, Yoo HJ, Fu SH, Hammock BD, Trainor BC: **Corncob bedding alters the effects of estrogens on aggressive behavior and reduces estrogen receptor-α expression in the brain**. *Endocrinology* 2011, **153**:949-953.

30. Sorge RE, Martin LJ, Isbester KA, Sotocinal SG, Rosen S, Tuttle AH, Wieskopf JS, Acland EL, Dokova A, Kadoura B *et al.*: **Olfactory exposure to males, including men, causes stress and related analgesia in rodents**. *Nat Methods* 2014, **11**:629-632.

31. Holder MK, Blaustein JD: **Puberty and adolescence as a time of vulnerability to stressors that alter neurobehavioral processes**. *Front Neuroendocrinol* 2014, **35**:89-110.

32. Kane L, Ismail N: **Puberty as a vulnerable period to the effects of immune challenges: focus on sex differences**. *Behav Brain Res* 2017, **320**:374-382.

33. Crabbe JC, Wahlsten D, Dudek BC: **Genetics of mouse behavior: interactions with laboratory environment**. *Science* 1999, **284**:1670-1672.

34. Collins LM, Dziak JJ, Kugler KC, Trail JB: **Factorial experiments**. *Am J Prev Med* 2014, **47**:498-504.

35. Collins LM, Dziak JJ, Li R: **Design of experiments with multiple independent variables: a resource management perspective on complete and reduced factorial designs**. *Psychol Methods* 2009, **14**:202-224.

36. Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR: **Power failure: why small sample size undermines the reliability of neuroscience**. *Nat Rev Neurosci* 2013, **14**:365-376.

37. Beery AK, Christensen JD, Lee NS, Blandino KL: **Specificity in sociality: mice and prairie voles exhibit different patterns of peer affiliation**. *Front Behav Neurosci* 2018, **12**.

38. Huang GZ, Woolley CS: **Estradiol acutely suppresses inhibition in the hippocampus through a sex-specific endocannabinoid and mGluR-dependent mechanism**. *Neuron* 2012, **74**:801-808.

39. Bales KL, Plotsky PM, Young LJ, Lim MM, Grotte N, Ferrer E, Carter CS: **Neonatal oxytocin manipulations have long-lasting, sexually dimorphic effects on vasopressin receptors**. *Neuroscience* 2007, **144**:38-45.

40. DeVries AC, DeVries MB, Taymans SE, Carter CS: **The effects of stress on social preferences are sexually dimorphic in prairie voles**. *Proc Natl Acad Sci* 1996, **93**:11980-11984.

41. Cho MM, DeVries AC, Williams JR, Carter CS: **The effects of oxytocin and vasopressin on partner preferences in male and female prairie voles (*Microtus ochrogaster*)**. *Behav Neurosci* 1999, **113**:1071.

42. Wu Z, Autry AE, Bergan JF, Watabe-Uchida M, Dulac CG: **Galanin neurons in the medial preoptic area govern parental behaviour**. *Nature* 2014, **509**:325-330.

43. Kohl J, Babayan BM, Rubinstein ND, Autry AE, Marin-Rodriguez B, Kapoor V, Miyamishi K, Zweifel LS, Luo L, Uchida N *et al.*: **Functional circuit architecture underlying parental behaviour**. *Nature* 2018, **556**:326-331.

44. Scott N, Prigge M, Yizhar O, Kimchi T: **A sexually dimorphic hypothalamic circuit controls maternal care and oxytocin secretion**. *Nature* 2015, **525**:519-522.

45. Zilkha N, Sofer Y, Beny Y, Kimchi T: **From classic ethology to modern neuroethology: overcoming the three biases in social behavior research**. *Curr Opin Neurobiol* 2016, **38**:96-108.

46. Yoest KE, Cummings JA, Becker JB: **Estradiol, dopamine and motivation**. *Cent Nerv Syst Agents Med Chem* 2014, **14**:83-89.

47. Jacobs E, D'Esposito M: **Estrogen shapes dopamine-dependent cognitive processes: implications for women's health**. *J Neurosci* 2011, **31**:5286-5293.