Report

Current Biology

# Sex Differences in Visual Motion Processing

## Highlights

- Males have shorter motion discrimination thresholds than females
- The difference between males and females replicates across three research sites
- Amplitude of MT+ responses predicts individual differences in motion discrimination
- Variability of MT+ responses, however, does not capture observed sex differences

## Authors

Scott O. Murray, Michael-Paul Schallmo, Tamar Kolodny, ..., Stefan J. Troche, Raphael A. Bernier, Duje Tadin

## Correspondence

somurray@uw.edu (S.O.M.), dtadin@ur.rochester.edu (D.T.)

## In Brief

Murray et al. find large and highly replicable sex differences in visual motion perception, highlighting the importance of sex as a factor in the design and analysis of perceptual and cognitive studies.

Murray et al., 2018, Current Biology *28*, 1–6
September 10, 2018 
https://doi.org/10.1016/j.cub.2018.06.014

CellPress

Please cite this article in press as: Murray et al., Sex Differences in Visual Motion Processing, Current Biology (2018), https://doi.org/10.1016/j.cub.2018.06.014

Current Biology
Report



# Sex Differences in Visual Motion Processing

Scott O. Murray,[1,7,*] Michael-Paul Schallmo,[2] Tamar Kolodny,[1] Rachel Millin,[1] Alex Kale,[1] Philipp Thomas,[3] Thomas H. Rammsayer,[3] Stefan J. Troche,[4] Raphael A. Bernier,[5] and Duje Tadin[6,*]
[1]Department of Psychology, University of Washington, Seattle WA 98195, USA
[2]Department of Psychiatry, University of Minnesota, Minneapolis, MN 55454, USA
[3]Department of Psychology, University of Bern, Bern, Switzerland
[4]Department of Psychology and Psychotherapy, University of Witten/Herdecke, Witten, Germany
[5]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA 98195, U.S.A
[6]Departments of Brain and Cognitive Sciences, Ophthalmology, and Neuroscience, University of Rochester, Rochester, NY 14627, USA
[7]Lead Contact
*Correspondence: somurray@uw.edu (S.O.M.), dtadin@ur.rochester.edu (D.T.)
https://doi.org/10.1016/j.cub.2018.06.014

## SUMMARY

**The importance of sex as a biological variable has recently been emphasized by major funding organizations [1] and within the neuroscience community [2]. Critical sex-based neural differences are indicated by, for example, conditions such as autism spectrum disorder (ASD) that have a strong sex bias with a higher prevalence among males [51, 3]. Motivated by this broader context, we report a marked sex difference in a visual motion perception task among neurotypical adults. Motion duration thresholds [4, 5]—the minimum duration needed to accurately perceive motion direction—were considerably shorter for males than females. We replicated this result across three laboratories and 263 total participants. This type of enhanced performance has previously been observed only in special populations including ASD, depression, and senescence [6–8]. The observed sex difference cannot be explained by general differences in speed of visual processing, overall visual discrimination abilities, or potential motor-related differences. We also show that while individual differences in motion duration thresholds are associated with differences in fMRI responsiveness of human MT+, surprisingly, MT+ response magnitudes did not differ between males and females. Thus, we reason that sex differences in motion perception are not captured by an MT+ fMRI measure that predicts within-sex individual differences in perception. Overall, these results show how sex differences can manifest unexpectedly, highlighting the importance of sex as a factor in the design and analysis of perceptual and cognitive studies.**

## RESULTS

We collected behavioral and neuroimaging data in the context of a larger research program examining neural mechanisms of visual perception in individuals with autism spectrum disorder (ASD). In this research, we also examined sex differences because of both a strong male bias in ASD prevalence [51, 3] and the broader relevance of sex as a biological variable [2, 9–12]. Here, we report a discovery we made, and subsequently replicated, by analyzing data from neurotypical participants. Of particular interest were measurements of motion duration thresholds—the minimum amount of time a stimulus needs to be displayed in order to accurately perceive its motion direction. Motion duration thresholds have been linked to motion processing in brain regions such as MT [13–15] and appear to be sensitive to fundamental neural computations such as surround suppression [5] and contrast gain control [15–17]. When these mechanisms are functioning properly, they cause an increase in the duration required to judge motion direction, especially for large stimulus sizes and high contrast levels [5, 15]. This notable feature of motion duration thresholds has motivated a broad range of studies with special populations [4], with examples of atypically shorter duration thresholds in ASD [6], old age [8], depression [7], and neurotypical subjects following transcranial magnetic stimulation (TMS)-induced disruption of MT function [18].

### Male versus Female Differences in Visual Motion Processing

We measured motion perception using an established experimental design [5, 6] that utilizes briefly presented drifting gratings. After each stimulus presentation, participants are asked to classify the stimulus as moving leftward or rightward. The presentation duration is adaptively adjusted on a trial-by-trial basis so that the individual finally achieves a predetermined level of performance (e.g., 80%). This two-alternative forced-choice approach minimizes criterion differences between individuals (e.g., responses being affected by decision biases and confidence [19]). Notably, participants are free to respond when ready (that is, we were not measuring reaction times). Thus, this non-speeded task measures how quickly participants can process visual motion independent of any differences in motor response speed. Analysis of performance in males and females revealed a large sex-related difference in overall duration thresholds. Male participants had significantly shorter duration thresholds than female participants (Figures 1A and 1B; a significant main effect of group in a mixed ANOVA; $F_{1,31} = 7.12$, $p = 0.01$). There was also a significant sex by contrast interaction ($F_{1,31} = 8.89$,

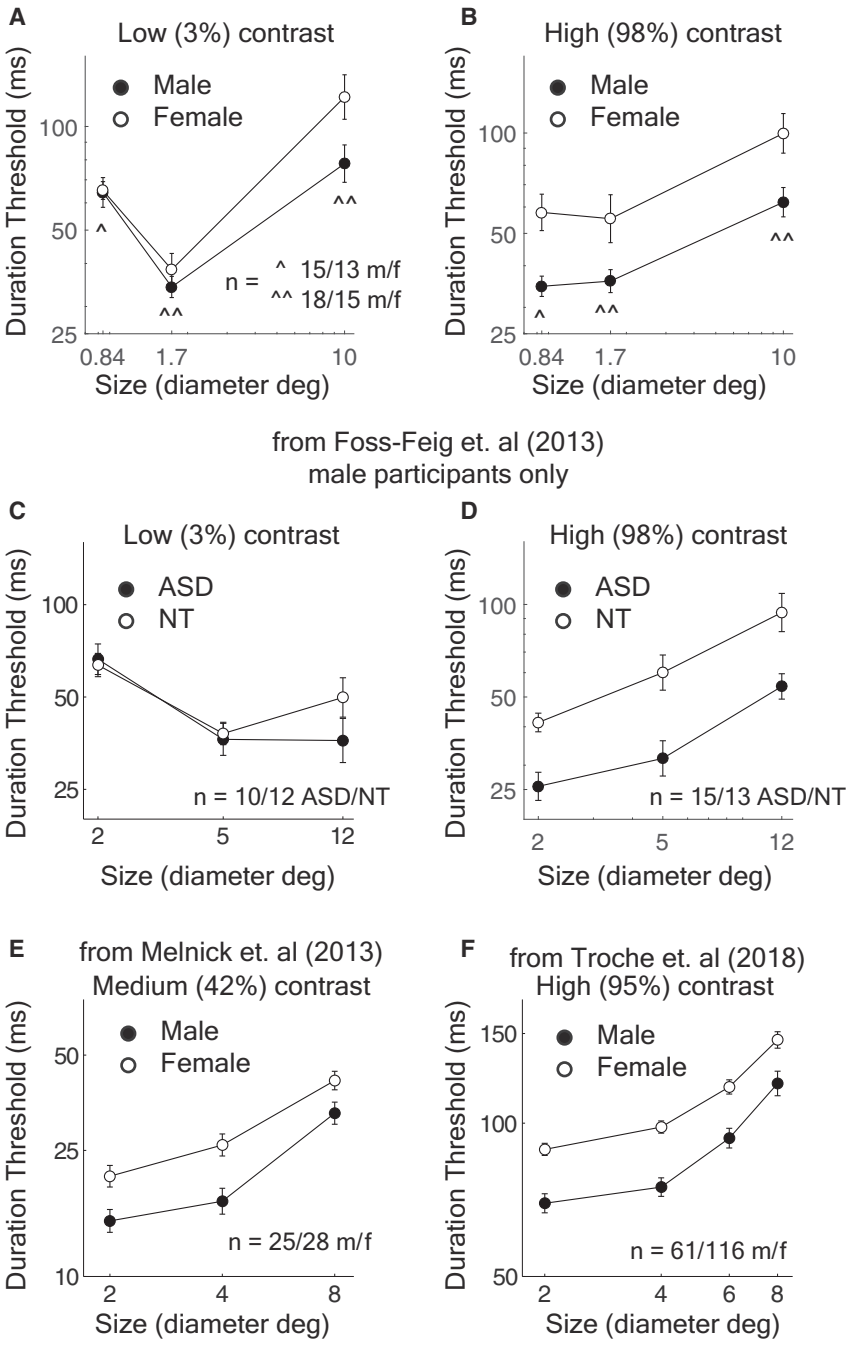


**Figure 1. Duration Thresholds**

(A–F) Males had shorter motion duration thresholds than females in Cohort 1 ([A] low-contrast performance; [B] high-contrast performance). This difference resembled previously reported findings in ASD [6] ([C] low-contrast performance; [D] high-contrast performance; five NT females participants were not included in these plots). The difference between males and females replicated across research sites and investigators ([E] Cohort 2; [F] Cohort 3) in groups with larger sample sizes. Error bars, SEM. See also Figures S1 and S2.

p = 0.001). Specifically, thresholds were 23% longer for females than males at low contrast and 78% longer at high contrast (expressed as a percentage to facilitate later comparisons across research sites). The effect size (Cohen's d) was 0.56 at low contrast and 1.11 at high contrast. As indicated by a significant three-way interaction ($F_{2,57}$ = 3.56, p = 0.035), at low contrast the effect of sex depended on stimulus size (Figure 1A).

This contrast-dependent pattern of results resembles those previously reported in ASD, where individuals with ASD showed shorter thresholds at high contrast compared to neurotypical individuals—a difference that largely disappeared at very low contrast [6]. To illustrate this qualitative similarity, we re-plotted the data from Foss-Feig et al. (2013) while including only the male participants (to avoid confounding the effects of ASD and gender; Figures 1C and 1D). This earlier study, which compared duration thresholds for ASD and neurotypical (NT) participants, reported a strikingly similar pattern to our sex-based differences, including a large difference at high contrast (Figure 1B versus 1D) and a size-dependent difference in thresholds for low contrast stimuli (Figure 1A versus 1C).

To examine the replicability of this surprising difference between males and females, we (the investigators of the study described above, which will be hereafter named "Cohort 1"; S.O.M., M.-P.S., T.K., R.M., A.K., R.A.B.) contacted other researchers (D.T., P.T., T.H.R., S.J.T.) with similarly designed studies and larger sample sizes. First, a previously published dataset [20] that used similar stimuli (contrast = 42%; sizes = 2°, 4°, 8° diameter) with a group of 53 individuals (25 male; 28 female) was reanalyzed ("Cohort 2"). The results revealed a similar sex-related difference (Figure 1E; a significant main effect of group in a mixed ANOVA; $F_{1,51}$ = 9.707, p = 0.003). The average increase in threshold for females in this previous study was 38% (Cohen's d = 0.86). Finally, in a recently published dataset [21] with 177 individuals (61 male; 116 female; "Cohort 3") and a similar design (contrast = 95%; sizes = 2°, 4°, 6°, 8° diameter), we once again observed this difference between males and females (Figure 1F; a significant main effect of group; $F_{1,175}$ = 23.3, $p < 10^{-5}$). The average difference in duration thresholds between males and females in this study was 27% (Cohen's d = 0.76). To further assess the reliability of the effect, we performed a meta-analysis that included the datasets from the three cohorts presented in Figures 1A, 1C, 1E, and 1F, as well as additional previous studies with smaller sample sizes than those above, and found a similar sex difference (see

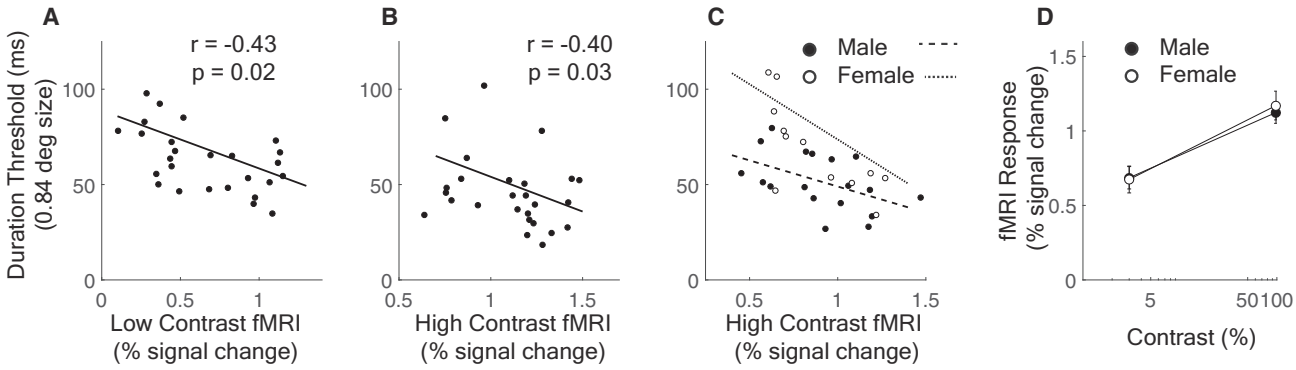


**Figure 2. MT+ fMRI Responses**

(A–D) Individual differences in MT+ amplitude were significantly correlated with individual differences in duration thresholds for (A) low- and (B) high-contrast stimuli. (C) The relationship between MT+ response and duration thresholds holds separately for males and females. (D) MT+ fMRI response for males and females is nearly identical. Error bars, SEM.

Figure S1). We note that while group plots convey a large and highly significant separation in performance between males and females (Figure 1), there is in fact considerable overlap in performance between the two populations as seen in distribution plots (see Figure S2).

As duration thresholds measure motion-processing speed, one possible explanation is that males are faster at rapidly processing visual information. To address this question, we examined additional data collected for Cohort 3. These same participants (n = 177) also completed a version of the Hick reaction time (RT) task [22]. There were four main conditions: a simple RT and choice RTs that required identifying the location of the target that could appear in two, four, or six different locations [21]. When considering choice RT conditions—conditions most similar to our motion task—we found no main effect of sex ($F_{1,175}$ = 1.856, p = 0.175) nor an interaction ($F_{2,350}$ = 0.001, p = 0.999). The results did not change when we considered all of the RT conditions (main effect of sex: $F_{1,175}$ = 2.840, p = 0.094; interaction: $F_{3,525}$ = 0.058, p = 0.982). Thus, we conclude that a general sex difference in visual processing speed cannot explain our main result. Further, we saw no difference in performance between males and females in a separate psychophysical task measuring contrast detection thresholds (see STAR Methods; Cohort 1: males, mean = 1.76%, SD = 0.93%; females, mean = 1.67%, SD = 0.66%; $t_{31}$ = 0.30, p = 0.77). There was also no difference between males and females in catch trial performance (see STAR Methods) during the motion discrimination task (Cohort 1: males, mean = 96% correct, SD = 0.05%; females, mean = 95% correct, SD = 0.06%; $t_{31}$ = 0.76, p = 0.45). In summary, our results suggest that the difference in motion discrimination between males and females may not be accounted for by general differences in task engagement as characterized by catch-trial performance, speed of visual processing, overall visual discrimination abilities, decision bias, or potential motor-related differences (since estimating duration threshold does not require speeded responses). We also examined whether the difference in motion discrimination between males and females might be accounted for by factors other than sex, such as age [8] or IQ [20]. These factors either did not differ or did not statistically contribute to the difference between males and females (see Supplemental Information).

### *Potential Neural Mechanisms*

To identify the potential underlying neural factors that may contribute to the performance differences between males and females, we first focused on the cortical area MT+. Motion duration thresholds have been hypothesized to reflect the response magnitude of neuronal subpopulations in area MT+ [4, 14, 15, 23]. Thus, we hypothesized that intrinsic, individual differences in neural responsiveness in MT+ would be associated with individual differences in duration thresholds. To test this prediction, we measured fMRI response amplitude in human MT+ in Cohort 1 in response to drifting sinewave gratings at two contrasts (3% and 98%) relative to a blank fixation condition. In order to emphasize sensory-driven responsiveness, participants' attention was diverted away from the motion stimuli with a central fixation task. Similar to previous studies [6], we focused on performance for the smallest stimulus size since any effects of spatial suppression and summation [5] on duration thresholds should be minimal. We found that individual differences in psychophysical thresholds correlated with individual differences in MT+ fMRI response magnitude for both low- (n = 28; $r_{26}$ = −0.43; p = 0.02; Figure 2A) and high-contrast stimuli (n = 28; $r_{26}$ = −0.40; p = 0.03; Figure 2B). An analogous pattern of results was found when we considered the relationship between MT+ responses averaged across contrast and duration thresholds averaged across size (medium and large; n = 33; $r_{31}$ = −0.45; p = 0.009). In addition, the relationship between individual differences in MT+ response magnitude and duration thresholds were present when males and females were analyzed separately (males, $r_{16}$ = −0.48, p = 0.04; females, $r_{13}$ = −0.55, p = 0.03; Figure 2C). Thus, motion duration threshold measurements appear to be a very good proxy for neural response magnitude in the key motion-processing area MT+. However, surprisingly, we found no significant sex differences in fMRI response amplitude in MT+ (Figure 2D; $F_{1,31}$ = 0.03, p = 0.86). Thus, the behavioral sex differences in motion perception are not captured by an MT+ fMRI measure that captures within-sex individual differences on the same task. We also investigated whether there were sex differences in response magnitudes in

early visual cortex (EVC; note that the foveal location of the stimuli prevented differentiating V1, V2, and V3 [24]), a source of cortical input for MT+, and found no differences ($F_{1,31} = 0.09$, $p = 0.76$).

## DISCUSSION

We report results that demonstrate a strong behavioral sex difference in visual motion perception. These results are notable given the increasing evidence supporting the gender similarities hypothesis—that, on average, men and women perform similarly on most cognitive tasks [25]. However, there are several apparent exceptions to the overall trend toward similar performance between males and females [26]. In meta-analyses, females were found to excel in verbal fluency ($d = -0.33$; [25]) and reading achievement ($d = -0.44$; [27]), while males tended to excel in measures of visual-spatial ability. Indeed, one of the largest and most documented sex differences in cognition is mental rotation [28], with the male performance advantage in the medium to large effect size range ($d \approx 0.50–0.80$, [29–31]). Thus, within the context of these previous findings, the sex differences in visual motion processing identified in the present study ($d = 0.56–1.11$) are at least as strong as the largest previously documented findings in other cognitive domains.

We also explored potential neural correlates of both individual and sex differences that were prominent in our behavioral data. The results showed that for both males and females, individual differences in task-independent, MT+ neural responsiveness was strongly associated with differences in psychophysical performance (Figure 2C). Thus, MT+ is likely a major contributor to the perception of briefly presented moving stimuli. However, while individual differences in responsiveness of human MT+ can explain individual differences in behavior, those same MT+ responses cannot explain the observed sex differences in behavior; response magnitudes in MT+ were nearly identical for males and females. There is no reason to suspect that our fMRI measurements were simply insensitive to underlying male and female differences in MT+ response magnitude. On the contrary, our fMRI measurements demonstrate a high degree of sensitivity both in the response to stimulus contrast (Figure 2D) and in correlations with behavior (Figures 2A–2C).

As our data appear to rule out arguably most obvious explanations, such as differences in MT+ responsiveness and visual-processing speed, our study largely leaves open the question of the underlying cause of observed sex differences. With respect to the role of MT+, our results cannot rule out the possibility that MT+ neurons are differentially engaged by males and females during task performance. Thus, future studies should investigate potential sex differences in MT+ responsiveness while attention is directed to the motion stimuli. Moreover, even for reliable sex differences such as mental rotation, other moderating variables such as sociocultural factors [32] as well as training and related experience [33, 34] can play significant roles. Of note, action video game playing results in large improvements in perceptual abilities [35–37], and males report more frequent video game play than females [38]. However, the specificity of sex differences to the motion task, and not to speed of processing and contrast discrimination tasks, is inconsistent with general improvements in visual function typically found to result from action video games [36, 39]. Moreover, we found sex differences in participants that completed the motion task with no prior task practice (Cohorts 1 and 3) and in those that completed a practice session administered a few days before the actual testing (Cohort 2). This appears to offer some differences between our results and sex differences in mental rotation, which are affected by training and action video game playing [33, 34].

Given the strong male prevalence of ASD, we find the observation that sex differences in motion perception closely resemble results reported for individuals with ASD [6] particularly interesting. Both neurotypical males (when compared to neurotypical females) and males with ASD (when compared to males without ASD) exhibit enhanced motion perception for high-contrast stimuli—a difference that is significantly attenuated at low contrast. The similarity of the sex difference and ASD difference is consistent with the hypothesis that ASD can be considered as an extreme version of the normal male profile [40]. However, the similarity between the sex differences and ASD differences should also be interpreted with caution given inconsistent findings among ASD studies on visual motion perception [41, 42].

The results from Cohort 1 demonstrating a strong contrast dependence in the sex difference are consistent with reduced gain control, a neural mechanism underlying the saturation of neural responses at high contrasts [43, 44]. Specifically, the brain typically suppresses responses to sensory stimuli when they become too intense—a mechanism that results in a relative reduction of responsiveness to high-contrast stimuli (saturation). One speculation is that males exhibit reduced contrast gain control and that this reduction is even more pronounced in males with ASD. However, we note that the lack of a difference in fMRI responsiveness (measured with attention directed away from the stimuli) suggests that differences in gain control between males and females may be tied to the deployment of attention [45] during the motion discrimination task.

In summary, our results clearly demonstrate a difference in motion processing between males and females and lend strong support to recent efforts to include sex as a biological factor in neuroscience research [2, 9–12]. In particular, our results have profound methodological implications for any research where there are inherent sex biases in subject populations. Overall, our results argue that any between-group sensory-perceptual experiments should either match the sex of participants or include it as a factor in the analyses.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBECT DETAILS
- METHOD DETAILS
  - fMRI experiment
  - Psychophysics experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental Information includes two figures and can be found with this article online at https://doi.org/10.1016/j.cub.2018.06.014.

## ACKNOWLEDGMENTS

The research at the University of Washington was supported by NIH R01MH106520 to S.O.M. and R.B.; M.-P.S. was supported by F32 EY025121. D.T.'s work was supported by NIH R01EY019295. We thank Randolph Blake for comments on an earlier draft of the manuscript and Jennifer Foss-Feig for helpful discussions related to these findings.

## AUTHOR CONTRIBUTIONS

S.O.M., M.-P.S., and R.A.B. designed the fMRI and psychophysics experiments at the University of Washington (U.W.) with Cohort 1. M.-P.S., T.K., R.M., and A.K. collected and analyzed (with S.O.M.) the Cohort 1 data at the U.W. D.T. conducted the re-analysis of the previous, Cohort 2, data at the University of Rochester. P.T., T.H.R., S.J.T., and D.T. conducted all aspects of the experiment at the University of Bern with Cohort 3. S.O.M., M.-P.S., and D.T. wrote the paper. All authors edited drafts of the manuscript.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 6, 2018
Revised: May 7, 2018
Accepted: June 8, 2018
Published: August 16, 2018

## REFERENCES

1. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-102.html
2. McCarthy, M.M., Woolley, C.S., and Arnold, A.P. (2017). Incorporating sex as a biological variable in neuroscience: what do we gain? Nature Reviews Neuroscience, nrn. 2017.2137.
3. Werling, D.M., and Geschwind, D.H. (2013). Sex differences in autism spectrum disorders. Curr. Opin. Neurol. *26*, 146–153.
4. Tadin, D. (2015). Suppressive mechanisms in visual motion processing: From perception to intelligence. Vision Res. *115* (Pt A), 58–70.
5. Tadin, D., Lappin, J.S., Gilroy, L.A., and Blake, R. (2003). Perceptual consequences of centre-surround antagonism in visual motion processing. Nature *424*, 312–315.
6. Foss-Feig, J.H., Tadin, D., Schauder, K.B., and Cascio, C.J. (2013). A substantial and unexpected enhancement of motion perception in autism. J. Neurosci. *33*, 8243–8249.
7. Golomb, J.D., McDavitt, J.R., Ruf, B.M., Chen, J.I., Saricicek, A., Maloney, K.H., Hu, J., Chun, M.M., and Bhagwagar, Z. (2009). Enhanced visual motion perception in major depressive disorder. J. Neurosci. *29*, 9072–9077.
8. Betts, L.R., Taylor, C.P., Sekuler, A.B., and Bennett, P.J. (2005). Aging reduces center-surround antagonism in visual motion processing. Neuron *45*, 361–366.
9. Bale, T.L., and Epperson, C.N. (2017). Sex as a biological variable: who, what, when, why, and how. Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology *42*, 386–396.
10. Joel, D., and McCarthy, M.M. (2017). Incorporating sex as a biological variable in neuropsychiatric research: where are we now and where should we be? Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology *42*, 379–385.
11. Miller, L.R., Marks, C., Becker, J.B., Hurn, P.D., Chen, W.-J., Woodruff, T., McCarthy, M.M., Sohrabji, F., Schiebinger, L., Wetherington, C.L., et al. (2017). Considering sex as a biological variable in preclinical research. FASEB J. *31*, 29–34.
12. Prager, E.M. (2017). Addressing sex as a biological variable. J. Neurosci. Res. *95*, 11.
13. Turkozer, H.B., Pamir, Z., and Boyaci, H. (2016). Contrast Affects fMRI Activity in Middle Temporal Cortex Related to Center-Surround Interaction in Motion Perception. Front. Psychol. *7*, 454.
14. Liu, L.D., Haefner, R.M., and Pack, C.C. (2016). A neural basis for the spatial suppression of visual motion perception. eLife *5*, e16167.
15. Schallmo, M.-P., Kale, A.M., Millin, R., Flevaris, A.V., Brkanac, Z., Edden, R.A., Bernier, R.A., and Murray, S.O. (2018). Suppression and facilitation of human neural responses. eLife *7*, e30334.
16. Betts, L.R., Sekuler, A.B., and Bennett, P.J. (2012). Spatial characteristics of motion-sensitive mechanisms change with age and stimulus spatial frequency. Vision Res. *53*, 1–14.
17. Rosenberg, A., Patterson, J.S., and Angelaki, D.E. (2015). A computational perspective on autism. Proc. Natl. Acad. Sci. USA *112*, 9158–9165.
18. Tadin, D., Silvanto, J., Pascual-Leone, A., and Battelli, L. (2011). Improved motion perception and impaired spatial suppression following disruption of cortical area MT/V5. J. Neurosci. *31*, 1279–1283.
19. Kingdom, F., and Prins, N. (2010). Psychophysics: a practical introduction (Academic Press London).
20. Melnick, M.D., Harrison, B.R., Park, S., Bennetto, L., and Tadin, D. (2013). A strong interactive link between sensory discriminations and intelligence. Curr. Biol. *23*, 1013–1017.
21. Troche, S.J., Thomas, P., Tadin, D., and Rammsayer, T.H. (2018). On the relationship between spatial suppression, speed of information processing, and psychometric intelligence. Intelligence *67*, 11–18.
22. Rammsayer, T.H., and Brandler, S. (2007). Performance on temporal information processing as an index of general intelligence. Intelligence *35*, 123–139.
23. Churan, J., Khawaja, F.A., Tsui, J.M., and Pack, C.C. (2008). Brief motion stimuli preferentially activate surround-suppressed neurons in macaque visual area MT. Curr. Biol. *18*, R1051–R1052.
24. Dougherty, R.F., Koch, V.M., Brewer, A.A., Fischer, B., Modersitzki, J., and Wandell, B.A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. J. Vis. *3*, 586–598.
25. Hyde, J.S. (2005). The gender similarities hypothesis. Am. Psychol. *60*, 581–592.
26. Hyde, J.S. (2014). Gender similarities and differences. Annu. Rev. Psychol. *65*, 373–398.
27. Reilly, D. (2012). Gender, culture, and sex-typed cognitive abilities. PLoS ONE *7*, e39904.
28. Voyer, D., Voyer, S.D., and Saint-Aubin, J. (2017). Sex differences in visual-spatial working memory: A meta-analysis. Psychon. Bull. Rev. *24*, 307–334.
29. Linn, M.C., and Petersen, A.C. (1985). Emergence and characterization of sex differences in spatial ability: a meta-analysis. Child Dev. *56*, 1479–1498.
30. Voyer, D. (2011). Time limits and gender differences on paper-and-pencil tests of mental rotation: a meta-analysis. Psychon. Bull. Rev. *18*, 267–277.
31. Voyer, D., Voyer, S., and Bryden, M.P. (1995). Magnitude of sex differences in spatial abilities: a meta-analysis and consideration of critical variables. Psychol. Bull. *117*, 250–270.
32. McGlone, M.S., and Aronson, J. (2006). Stereotype threat, identity salience, and spatial reasoning. J. Appl. Dev. Psychol. *27*, 486–493.
33. Feng, J., Spence, I., and Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. Psychol. Sci. *18*, 850–855.
34. Uttal, D.H., Meadow, N.G., Tipton, E., Hand, L.L., Alden, A.R., Warren, C., and Newcombe, N.S. (2013). The malleability of spatial skills: a meta-analysis of training studies. Psychol. Bull. *139*, 352–402.
35. Green, C.S., and Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. Psychol. Sci. *18*, 88–94.

36. Li, R., Polat, U., Makous, W., and Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. Nat. Neurosci. *12*, 549–551.
37. Bejjanki, V.R., Zhang, R., Li, R., Pouget, A., Green, C.S., Lu, Z.-L., and Bavelier, D. (2014). Action video game play facilitates the development of better perceptual templates. Proc. Natl. Acad. Sci. USA *111*, 16961–16966.
38. Lucas, K., and Sherry, J.L. (2004). Sex differences in video game play: A communication-based explanation. Communic. Res. *31*, 499–523.
39. Dye, M.W., Green, C.S., and Bavelier, D. (2009). Increasing speed of processing with action video games. Curr. Dir. Psychol. Sci. *18*, 321–326.
40. Baron-Cohen, S. (2002). The extreme male brain theory of autism. Trends Cogn. Sci. *6*, 248–254.
41. Schauder, K.B., Park, W.J., Tadin, D., and Bennetto, L. (2017). Larger Receptive Field Size as a Mechanism Underlying Atypical Motion Perception in Autism Spectrum Disorder. Clin. Psychol. Sci. *5*, 827–842.
42. Sysoeva, O.V., Galuta, I.A., Davletshina, M.S., Orekhova, E.V., and Stroganova, T.A. (2017). Abnormal Size-Dependent Modulation of Motion Perception in Children with Autism Spectrum Disorder (ASD). Front. Neurosci. *11*, 164.
43. Geisler, W.S., and Albrecht, D.G. (1992). Cortical neurons: isolation of contrast gain control. Vision Res. *32*, 1409–1410.
44. Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. Vis. Neurosci. *9*, 181–197.
45. Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. Neuron *61*, 168–185.
46. Brainard, D.H. (1997). The Psychophysics Toolbox. Spat. Vis. *10*, 433–436.
47. Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. *10*, 437–442.
48. Kleiner, M., Brainard, D.H., and Pelli, D.G. (2007). What's new in Psychtoolbox 3. in Perception 36 ECVP Abstract Supplement.
49. Prins, N., and Kingdom, F.A.A. (2009). Palamedes: MATLAB routines for analyzing psychophysical data.
50. Watson, A.B., and Pelli, D.G. (1983). QUEST: a Bayesian adaptive psychometric method. Percept. Psychophys. *33*, 113–120.
51. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators (2014). Prevalence of autism spectrum disorder among children aged 8 years – autism and developmental disabilities monitoring network, 11 sites, United States, 2010. Morbidity and Mortality Weekly Report *63*, 1–21.

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Software and Algorithms | | |
| Brainvoyager QX | Brain Innovation | 2.8.4.2645 |
| MATLAB | Mathworks | R2013b |

## CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Scott O. Murray (somurray@uw.edu).

## EXPERIMENTAL MODEL AND SUBECT DETAILS

There were three separate cohorts of participants (total n = 263), each tested at a different research site that included different investigators. Cohort 1 (n = 33) included 18 males and 15 females from 18 to 30 years of age who participated in psychophysical experiments and fMRI experiments at the University of Washington, U.S.A with investigators S.O.M., M-P.S., T.K., R.M., A.K., and R.A.B. Cohort 2 (n = 53) included 25 males and 28 females from 18 to 59 years of age who participated in psychophysical experiments at the University of Rochester, U.S.A. with investigator D.T. Cohort 3 (n = 177) included 61 males and 116 females ranging in age from 18 to 30 that participated in psychophysical experiments at the University of Bern, Switzerland with investigators P.T., T.H.R., S.J.T., and D.T.. All participants across all three cohorts had normal or corrected-to-normal vision. For all experiments, sex was determined by a binary self-report of male or female. All procedures were approved by the respective Institutional Review Boards and all participants provided written informed consent.

## METHOD DETAILS

### fMRI experiment

Data were acquired using a Philips Achieva 3 Tesla scanner. T1-weighted structural MRI data were acquired in each session at 1 mm isotropic resolution. Functional MRI data (gradient echo EPI) were acquired with 3 × 3 mm inplane resolution. 30 oblique-axial slices were obtained (3 mm slice thickness, separated by a 0.5 mm gap). Other scan parameters: 2 s TR, 25 ms TE, 79° flip angle, A-P phase-encode direction. At the start of each session, an opposite phase-encode direction (P-A) scan (1 TR) was acquired for distortion compensation. Each scanning session lasted about 1 hour.

Stimuli were displayed via projector (Epson Powerlite 7250 or Eiki LCXL100A, following a hardware failure), operating at 60 Hz using Presentation software (Neurobehavioral Systems, Berkeley, CA) on a PC running Windows XP. Images were projected on a semicircular screen at the rear of the scanner, and viewed through a mirror on the head coil at a distance of 66 cm. Projector luminance was linearized using custom software.

The experiment was designed to measure the amplitude of the response in human MT+ and early visual cortex (EVC) to two stimulus contrast levels. Sixteen gratings were presented at the center of the screen during each block; to prevent adaptation, gratings moved in one of eight possible directions in a randomized and counter-balanced order for 400ms. We used longer stimulus durations than typical psychophysical duration thresholds to ensure robust BOLD responses to moving stimuli. Twenty-five blocks were presented during each run (10 s each, 125 TRs total). Stimulus diameter was 2°, and contrast varied across blocks. Each run began with a blank block (0% contrast), in which only the fixation task was presented on a blank background. Blocks of drifting gratings at 3% and 98% contrast were then presented centrally in an alternating order, each followed by a blank block (6 low contrast, 6 high contrast, and 13 blank blocks per run). The blank block served as baseline. A 10 s baseline period was chosen to balance between the competing needs of accurately estimating response amplitudes – for which a long (20-30 s) baseline period would afford full recovery of the hemodynamic response – and increasing the number of trials in the experiment to maximize signal-to-noise. To functionally localize MT+ we used a separate localizer scan. Drifting and static 2° gratings (15% contrast) were presented centrally in alternating 10 s blocks (13 static and 12 drifting blocks, 125 TRs total). Another localizer scan was used to functionally identify EVC. Here, contrast-reversing checkerboard stimuli (2° diameter, 100% contrast, 8 Hz) were presented at the center of the screen, and alternated with blank backgrounds across a total of 16 blocks (8 stimulus & blank blocks, 10 s per block, 80 TRs total). Subjects completed two runs per session and all but one subject completed two fMRI sessions (other experiments, not presented here, were included in the fMRI sessions).

During all scans, subjects performed a color & shape conjunction detection task, responding to a green circle within a series of colored shapes presented briefly at fixation (Figure 2A). A total of 8 shapes were presented in each 10 s block, and were selected randomly from 8 possible shapes (including the green circle target). Hit rate in this task (averaged across all scans within each subject) did not differ between males (mean = 95%, SD = 6.7%) and females (mean = 95%, SD = 7.6%; $t_{24} = 0.01$, $p = 0.99$).

FMRI data were processed in BrainVoyager (Brain Innovation, Maastricht, the Netherlands). The following steps were performed in order: motion correction, distortion compensation, high-pass filtering (cutoff = 2 cycles/scan), and alignment to the T1 anatomy. No spatial smoothing or normalization were performed for ROI-based, within-subjects analyses. ROIs were identified from the functional localizer data using correlational analyses, with an initial significance threshold of $p < 0.05$ (Bonferroni corrected). The top 20 most-significant voxels within the ROI were then selected for further analyses. In cases where there were not 20 voxels that satisfied the above threshold, the threshold was relaxed until 20 voxels were included. ROIs were defined for each hemisphere in 2 anatomical regions: motion-selective MT+ in the lateral occipital lobe, and the region of EVC representing the center stimulus (near the occipital pole). ROI position was verified by visualization on an inflated cortical surface.

Average fMRI time courses were extracted from each ROI for further analyses in MATLAB using BVQXTools. Time course data were divided into epochs from 4 s before to 4 s after each block. For each block, response baseline was calculated as the average signal across all epochs between 0-4 s prior to block onset. The time course in each block was then converted to percent signal change. Time courses were then averaged across hemispheres in each run, and across runs in each subject. The response peak, defined as the average signal change from 8-12 s after the block onset, served as the measure of the fMRI response. A full-brain, voxel-wise mixed ANOVA (sex X contrast) was also performed on spatially (Talairach) and temporally (%-transform) normalized data.

## Psychophysics experiments

For all three research sites, psychophysical experiments were similar. All involved visual motion direction discriminations of time-limited grating-based stimuli. Temporal envelopes were trapezoidal, with the duration defined by the full-width at half-maximum contrast of the temporal envelope. As detailed below, however, there were a number of apparatus, stimulus and procedural differences. For example, there are differences in stimulus contrast, stimulus size, spatial stimulus envelope, visual display type, and threshold estimation procedures. Given that similar sex differences were observed at all three research sites, we can conclude that the observed sex differences generalize over these variations in stimulus parameters and task procedures.

### *Cohort 1 at the University of Washington*

Stimuli were presented using a ViewSonic PF790 CRT monitor (120 Hz) with an associated Bits# stimulus processor (Cambridge Research Systems, Kent, UK). The monitor luminance was linearized using custom software. Stimuli were presented on a Windows PC in MATLAB (MathWorks, Natick, MA) using Psychtoolbox-3 [46-48], with a chin rest used to stabilize head position at a viewing distance of 66 cm.

In each experiment, we presented drifting sinusoidal luminance modulation gratings at two different Michelson contrast levels (low = 3%, high = 98%) and 3 different sizes (diameter = 0.84, 1.7 and 10°). Motion speed was 4 cycles/s, and spatial frequency was 1.2 cycles/°. Gratings were presented within a circular aperture, whose edges were blurred with a Gaussian envelope (SD = 0.21°). Stimuli were presented centrally on a mean luminance background.

Trials began with a central fixation mark, a small shrinking circle (850 ms). This was followed by a blank screen (150 ms), after which the grating stimuli appeared (variable duration controlled by a staircase procedure, range 6.7 – 333 ms), followed by another blank screen (150 ms), and finally a fixation mark (the response cue). Subjects indicated their response (left or right) using the corresponding arrow keys. Response time was not limited. To permit very brief stimulus presentations, gratings appeared within a trapezoidal temporal envelope, according to an established method [6]. Duration of the grating stimuli varied across trials according to a Psi adaptive staircase procedure [19] controlled using the Palamedes toolbox [49]. Duration was adjusted across subsequent trials based on task performance, to determine the amount of time needed to correctly discriminate motion direction with 80% accuracy (i.e., the threshold duration). Feedback was not provided. Staircases were run separately to determine thresholds for each of the six stimulus conditions (2 contrasts x 3 sizes, as above). Condition order was randomized across trials. Thirty trials were collected per staircase within a single run (approximately 6 min). There were also 10 catch trials per run (all 10° diameter, 98% contrast gratings, 333 ms duration), which were used to assess off-task performance. Data from one female was excluded from this cohort (i.e., not included in the 15 females) due to poor performance on the catch trials (accuracy < 80%).

Each subject completed 4 runs in the motion discrimination paradigm, with a total experiment duration of about 30 min. Example and practice trials were presented before the first experiment run. For 5 subjects, thresholds were not obtained for the smallest stimulus size. Psychometric thresholds and slopes were quantified for each condition in each run by fitting the discrimination accuracy data with a Weibull function using maximum likelihood estimation [19]. Guess rate and lapse rate were fixed at 50% and 4%, respectively. Threshold duration was defined at 80% accuracy based on this fit. Average thresholds for each subject were found by taking the median value across the 4 runs.

Subjects in Cohort 1 also completed a contrast detection paradigm. The display apparatus was the same as in the motion paradigm. Using the Bits# stimulus processor in mono++ mode allowed us to present stimuli with 9.6 bit luminance resolution, which permitted the presentation of very low contrast stimuli. The task was to detect whether a Gabor (sinusoidal luminance modulation within a Gaussian window, SD = 0.42°, FWHM = 1°, 1.5 cycles/°, vertical or horizontal orientation) was presented at the center of a mean gray background during either the first or second of two possible stimulus presentation intervals. Trial order was as follows. First, a '1' was presented at fixation for 350 ms (to denote the first presentation interval). Then, the same fixation mark as above

(shrinking circle) was presented, followed by either the Gabor or a blank background. Gabor (and blank) duration was 100 ms, and there was a 1 s blank between intervals. Then, a '2' was presented (as above), followed by the fixation circle, and then the second stimulus presentation interval (blank or Gabor). Presentation of the Gabor within the first or second interval was randomized and counter-balanced across trials. Stimulus contrast was adjusted using the same staircase procedure as above, to determine the lowest contrast that could be detected with 80% accuracy. Vertically and horizontally-oriented Gabors were presented in separate, interleaved staircases (order randomized across trials). There were 30 trials per staircase in each run, as well as 20 catch trials (contrast fixed at 45%). Subjects completed 2 runs of the contrast detection paradigm (total duration about 15 min). The mean threshold from both conditions in both staircases was used for analysis.

#### *Cohort 2 at the University of Rochester*

Only key methodological details are reported here. For detailed methods see [20]. Stimuli were created and displayed in MATLAB and shown on a custom 360Hz DLP projector (natively linear, 1280 × 720 resolution, 113.7 cd/m2 background) with a viewing distance of 146 cm. Stimuli were briefly presented sine-wave gratings, moving either leftward or rightward with speed of 4°/s. Peak stimulus contrast was 42%, Stimulus spatial frequency was 1 cycle/°. Stimulus duration was defined as the full-width at half-height of the trapezoidal temporal envelope (as detailed in [18]). Stimulus size was defined by the stationary raised cosine spatial envelope through which moving gratings were shown. There were three different stimulus sizes (2°, 4°, and 8° in diameter). Different sizes were interleaved within each block.

On each trial, a moving stimulus was presented and a subject identified perceived motion direction (leftward versus rightward). Feedback was provided. The dependent variable was log10(stimulus duration). For each condition, eight duration thresholds (82% correct) were estimated by QUEST staircases [50]. The first two were measured on a separate day and used as practice. For each participant, the highest and lowest staircase results were excluded and the remaining thresholds were averaged. The reported sex difference remains similarly significant even if the extreme results were not excluded.

#### *Cohort 3 at the University of Bern*

Detailed methods appear in the original publication of these data [21]. Briefly, a linearized LCD monitor (Asus VG248Qe, 144Hz, 1920 × 1080 resolution) was used with a viewing distance of 61 cm that was maintained with a chin rest. Stimulus presentation was controlled with MATLAB. Stimuli consisted of 95% contrast gratings with a spatial frequency of 1 cycle/° moving at 4.8°/s. Stimulus size was determined by stationary raised cosine spatial envelopes. Expressed as the full stimulus diameter, stimulus sizes were 2°, 4°, 6° and 8°. (In a previous publication [21], reported diameters were smaller because they were based on stimulus values above 1% contrast). After a practice session of 180 trials, a total of three blocks were administered, with each block consisting of 44 trials of each stimulus size resulting in 528 trials. The four stimulus sizes were randomly interleaved within each block. On each trial, a moving stimulus was presented in the center of the participant's visual field. After the presentation of the stimulus, the participant indicated the perceived motion direction of the drifting grating (either leftward or rightward). Each correct response was followed by auditory feedback. For each stimulus size, six estimates of the 82%-detection threshold for motion perception were obtained using a Bayesian adaptive QUEST procedure [50] to estimate the presentation time required by a given participant to produce 82% correct responses. The highest and the lowest threshold estimates for each stimulus size were excluded. For each participant, the remaining four threshold estimates were averaged separately for each stimulus size.

## QUANTIFICATION AND STATISTICAL ANALYSIS

Pearson correlation was used to assess the relationship between individual differences in fMRI response magnitude and psychophysically measured duration threshold. Mixed between (low/high MT-response and male/female), within (stimulus-size) ANOVAs were used to assess between-group differences in psychophysical performance (Figures 1) and differences in fMRI response (Figure 2).

**Current Biology, Volume *28***

## Supplemental Information

# Sex Differences in Visual Motion Processing

**Scott O. Murray, Michael-Paul Schallmo, Tamar Kolodny, Rachel Millin, Alex Kale, Philipp Thomas, Thomas H. Rammsayer, Stefan J. Troche, Raphael A. Bernier, and Duje Tadin**

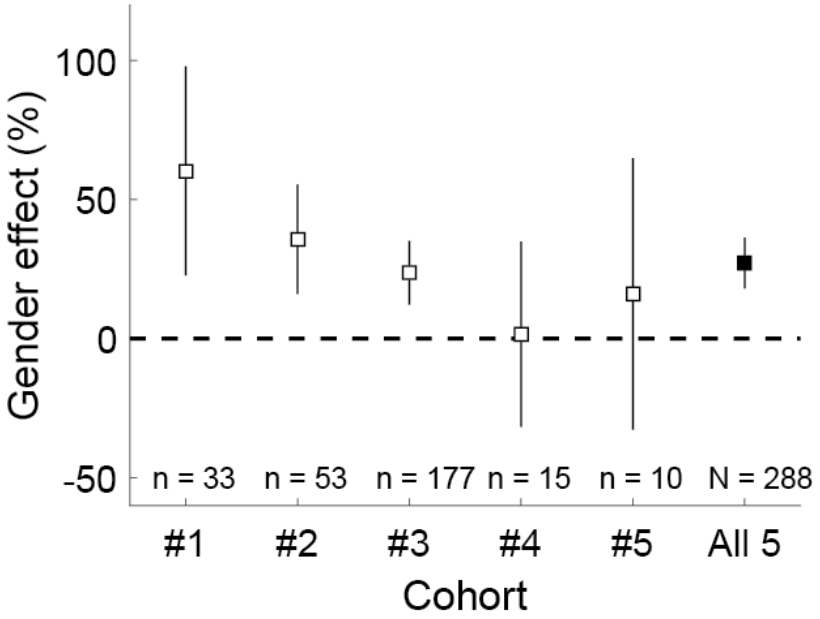


**Figure S1. Meta-analysis of sex effect on motion discrimination. Related to Figure 1.** Points show the mean difference in thresholds between females and males for each cohort (as a percentage of the mean male threshold). Error bars are 95% CIs. The number of subjects is reported along the x-axis for each cohort, as well as for the pooled data from all 5 cohorts (on the right).

A meta-analysis was performed to determine the reliability and magnitude of the sex effect across different motion discrimination experiments. Five data sets were included: the first 3 are described in the main Methods text, and data are presented in Figure 1. Cohorts 4 & 5 were part of a recent study from our group at the University of Washington [S1]. A significant sex effect was found across the 5 cohorts (ANOVA main effect, $F_{1,282} = 19.1$, $p = 2 \times 10^{-5}$).

Next, data were pooled across the 5 cohorts after converting to a percentage (relative to the mean of all males in that data set). The average sex effect pooled across cohorts, and 95% confidence intervals (CIs) for the pooled data were then calculated. We found a 27.2% increase in thresholds for females versus males across the 5 cohorts (CI: 18.1 – 36.3%). These data are presented in Supplemental Figure 1, and illustrate the magnitude and reliability of the observed sex effect.

Details of the experimental paradigms for data sets 4 and 5 are provided elsewhere [S1] – briefly, the paradigm in data set 4 was identical to the paradigm used in experiment 1 of the current study, while the paradigm in data set 5 used slightly larger stimuli (diameters of 1, 2, and 12°). Data set 4 is from the placebo session in the "lorazepam experiment" in the previous study, while data set 5 is from the "initial psychophysics experiment." Note that 4 subjects (all males) were included in both Cohorts 4 and 5 – we chose to retain these subjects in both samples as their data were collected in separate experimental sessions on different days using slightly different paradigms and display apparatuses.

The significance of the sex effect across cohorts was assessed using a mixed ANOVA. Subjects and cohorts were both treated as random effects, such that the results of this statistical test may generalize outside of this particular sample of subjects and cohorts. Stimulus size was treated as a continuous variable. Stimulus contrast was not considered in this analysis, as Cohorts 2 and 3 did not include data from multiple contrast levels.

The magnitude of the sex effect was also assessed. In each data set, thresholds were normalized to the average of all males from that cohort. We again collapsed across stimulus contrasts (for Cohorts 1, 4, and 5) in this analysis. An average sex effect (females minus males) for each dataset was then taken, and confidence intervals were calculated based on the pooled variance for males and females [S2]. Average sex effect for all data sets pooled together and CIs were also calculated (Figure S1).

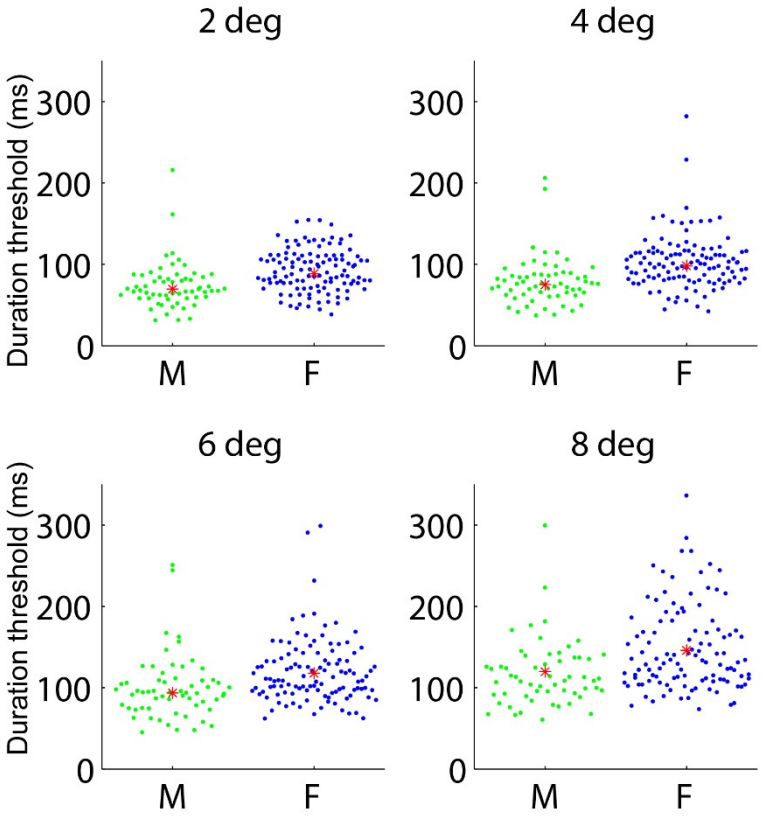


**Figure S2. Distribution of duration thresholds. Related to Figure 1.** Duration thresholds for individuals from Cohort 3 for each stimulus size. Each dot represents an individual subject. M = male, F = female. Red asterisks are the mean of the distribution.

**Age and IQ Analyses**

Average age was slightly higher for females in Cohort 1 (males mean = 22.3 years, SD = 2.4; females mean = 24.9 years, SD = 4.1; $t_{31} = 2.2$, $p = 0.03$). There was no link between age and duration thresholds ($r_{31} = 0.17$, $p = 0.34$). IQ did not significantly differ between males and females in Cohort 1 (males mean = 112.8, SD = 14.7; females mean = 113.5, SD = 11.4; $t_{31} = 0.15$, $p = 0.88$). Age (males mean = 31.1 years, SD = 12.6; females mean = 35.0 years, SD = 14.0; $t_{51} = 1.08$, $p = 0.28$) and IQ (males mean = 114, SD = 14.6; females mean = 112, SD = 11; $t_{51} = 0.84$, $p = 0.41$) did not differ in Cohort 2. For Cohort 3, males were slightly older than females (males mean = 21.9 years, SD = 3.5; females mean = 20.8 years, SD = 2.1; $t_{175} = 2.63$, $p = 0.01$). As with Cohort 1, there was no link between age and duration thresholds ($r_{175} = 0.05$, $p = 0.48$). As with the other cohorts, there were no significant sex differences in intelligence test performance in Cohort 3 ($t_{175} = 1.505$, $p = 0.134$). In addition, ANCOVAs were conducted to determine if there was a statistically significant difference between males and females on duration thresholds controlling for age and IQ. The results revealed similar sex-related differences with a significant main effect of sex in Cohort 1 ($F_{1,24} = 11.03$, $p = 0.003$), Cohort 2 ($F_{1,49} = 7.683$, $p = 0.008$), and Cohort 3 ($F_{1,173} = 31.701$, $p < 0.0001$).

**Supplemental References**

[S1] Schallmo, M.-P., Kale, A.M., Millin, R., Flevaris, A.V., Brkanac, Z., Edden, R.A., Bernier, R.A., and Murray, S. (2018). Suppression and facilitation of human neural responses. eLife *7*, e30334.

[S2] Howell, D. C. (2011). Confidence intervals on effect size. Retrieved August 11, 2017, from http://www.uvm.edu/~dhowell/methods8/ Supplements/Confidence%20Intervals%20on%20Effect%20Size.pdf.