Sex beyond the genitalia: The human brain mosaic

Daphna Joel\textsuperscript{a,b,1}, Zohar Berman\textsuperscript{b}, Ido Tavor\textsuperscript{c}, Nadav Wexler\textsuperscript{d}, Olga Gaber\textsuperscript{e}, Yaniv Stein\textsuperscript{f}, Nisan Shefi\textsuperscript{ab}, Jared Pool\textsuperscript{e}, Sebastian Ursche\textsuperscript{g}, Daniel S. Margulies\textsuperscript{h}, Franziskus Liem\textsuperscript{b,i}, Jürgen Hänggi\textsuperscript{f}, Lutz Jäncke\textsuperscript{d}, and Yaniv Assaf\textsuperscript{b,c}

\textsuperscript{a}School of Psychological Sciences, Tel-Aviv University, Ramat Aviv, Tel-Aviv 6997801, Israel; \textsuperscript{b}Sagol School of Neurosciences, Tel-Aviv University, Ramat Aviv, Tel-Aviv 6997801, Israel; \textsuperscript{c}Department of Neurobiology, Faculty of Life Sciences, Tel-Aviv University, Ramat Aviv, Tel-Aviv 6997801, Israel; \textsuperscript{d}School of Mathematical Sciences, Tel-Aviv University, Ramat Aviv, Tel-Aviv 6997801, Israel; \textsuperscript{e}Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany; and \textsuperscript{f}Division Neuropsychology, Department of Psychology, University of Zurich, 8092 Zurich, Switzerland

Edited by Bruce S. McEwen, The Rockefeller University, New York, NY, and approved October 23, 2015 (received for review June 4, 2015)

Whereas a categorical difference in the genitalia has always been acknowledged, the question of how far these categories extend into human biology is still not resolved. Documented sex/gender differences in the brain are often taken as support of a sexually dimorphic view of human brains ("female brain" or "male brain"). However, such a distinction would be possible only if sex/gender differences in brain features were highly dimorphic (i.e., little overlap between the forms of these features in males and females) and internally consistent (i.e., a brain has only "male" or only "female" features). Here, analysis of MRIs of more than 1,400 human brains from four datasets reveals extensive overlap between the distributions of females and males for all gray matter, white matter, and connections assessed. Moreover, analyses of internal consistency reveal that brains with features that are consistently at one end of the "maleness-femaleness" continuum are rare. Rather, most brains are comprised of unique "mosaics" of features, some more common in females compared with males, some more common in males compared with females, and some common in both females and males. Our findings are robust across sample, age, type of MRI, and method of analysis. These findings are corroborated by a similar analysis of personality traits, attitudes, interests, and behaviors of more than 5,000 individuals, which reveals that internal consistency is extremely rare. Our study demonstrates that, although there are sex/gender differences in the brain, human brains do not belong to one of two distinct categories: male brain/female brain.

The question of whether males and females form two distinct categories has attracted thinkers from ancient times to this day. Whereas a categorical difference in the genitals has always been acknowledged, the question of how far these categories extend into human biology is still not resolved (for a historical overview, see refs. 1 and 2). Documented sex/gender differences in the brain are often taken as support of a sexually dimorphic view of human brains ("female brain" vs. "male brain"), and consequently, of a sexually dimorphic view of human behavior, cognition, personality, attitudes, and other gender characteristics (3). Joel (4, 5) has recently argued that the existence of sex/gender differences in the brain is not sufficient to conclude that human brains belong to two distinct categories. Rather, such a distinction requires the fulfillment of two conditions: one, the form of the elements that show sex/gender differences should be dimorphic, that is, with little overlap between the forms of the elements in males and females. Two, there should be a high degree of internal consistency in the form of the different elements of a single brain (e.g., all elements have the "male" form). Previous criticisms of the dichotomous view of human brain have focused on the fact that most sex/gender differences are non-dimorphic population-level differences with extensive overlap of the distributions of females and males and have therefore claimed that human brains cannot be sorted into two distinct classes: "male brains" and "female brains" (6–8). However, if brains are internally consistent in the degree of "maleness-femaleness" of each of their elements, it will still be possible to align brains on a "male-brain–female-brain" continuum (4, 5). Such an alignment may be predicted by the clear view of sexual differentiation of the brain, according to which masculinization and feminization of the brain are under the sole influence of testosterone (9). In contrast, more recent evidence that masculinization and feminization are independent processes and that sexual differentiation progresses independently in different brain tissues (10), predicts poor internal consistency (4, 5). Poor internal consistency is further predicted by evidence that the effects of sex may be different and even opposite under different environmental conditions and that these sex-by-environment interactions may be different for different brain features (4, 5).

There are indeed examples of lack of internal consistency within a single brain in the animal literature (4, 5), yet it is not clear whether this is a common phenomenon that involves most features that show sex differences and is seen in most individuals. Here we assess the degree of internal consistency in the human brain using data obtained from MRI, a method that allows the simultaneous assessment of multiple brain features in many individuals.

We used datasets obtained from several different imaging modalities and analyzed with different methods to ensure that our conclusion is not measure, analysis, or sample dependent.

Significance

Sex/gender differences in the brain are of high social interest because their presence is typically assumed to prove that humans belong to two distinct categories not only in terms of their genitalia, and thus justify differential treatment of males and females. Here we show that, although there are sex/gender differences in brain and behavior, humans and human brains are comprised of unique "mosaics" of features, some more common in females compared with males, some more common in males compared with females, and some common in both females and males. Our results demonstrate that regardless of the cause of observed sex/gender differences in brain and behavior (nature or nurture), human brains cannot be categorized into two distinct classes: male brain/female brain.

Author contributions: D.J. designed research; D.J., Z.B., S.U., F.L., J.H., and J.L. performed research; D.J. and Y.A. contributed new analytic tools; D.J., Z.B., I.T., N.W., O.G., Y.S., N.S., J.P., S.U., F.L., and J.H. analyzed data; and D.J. and D.S.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Freely available online through the PNAS open access option.

Data deposition: Our anonymized raw neuroimaging data and accompanying metadata have been deposited at psy-neuro-nassy.uzh.ch and are accessible with a username and password that can be obtained from the authors by email (djioel@post.tau.ac.il or j.haengg@psychologie.uzh.ch).

1To whom correspondence should be addressed. Email: djioel@post.tau.ac.il.

This article contains supporting information online at www.pnas.orglookupupplp doi:10.1073/pnas.1509654112/2CSupplemental.

We use the term sex/gender to indicate that studies typically assess subjects’ sex (i.e., whether one is male or female) but observed differences may reflect the effects of both sex and gender (that is, the social construction of sex). We ignore here the important issue of the probable effects of gender on observed differences between females and males in brain and behavior, because we want to emphasize that regardless of the cause of these differences (sex, gender, or their interactions), they do not add up to create two distinct categories, one typical of males and the other typical of females.
The number of subjects in these datasets ranged from 138 to 855. In each dataset, following an assessment of sex/gender differences in all regions, we focused on the regions showing the largest sex/gender differences (i.e., least overlap between females and males). Because also in these regions there was a considerable overlap between the distributions of females and males, which made a division into two distinct forms impossible, we tested whether individuals would be consistently at one end of the “femaleness-maleness” continuum across brain regions or show “substantial variability”, being at the one end of the “femaleness-maleness” continuum on some regions and at the other end on other regions. We found that regardless of sample, type of MRI, and method of analysis, substantial variability is much more prevalent than internal consistency.

Table 1. Internal consistency and substantial variability in human brain and behavior

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Age: range, mean (SD)</th>
<th>Number of characteristics assessed for sex/gender differences</th>
<th>Percent of brains/individuals with substantial variability: both male-end and female-end features</th>
<th>Percent of brains/individuals with internal consistency: only female-end (F), only intermediate (I), or only male-end (M) (all, c, φ)</th>
<th>Average (SD) percent of features at the female-end (F) and male-end (M) zones (c, φ) and Cohen’s $d$ of the sex/gender difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>First sample, VBM</td>
<td>c: 18–79, 31.5 (12.0)</td>
<td>10 (116) 0.70 &lt; $d$ ≤ 0.84 all P &lt; 0.0001 c: 35% I: 3.6, 3.6, 3.6% M: 2.0, 3.0, 0.0% F: 2.2, 0.0, 4.3%</td>
<td>All: 35% F: 0.4, 0.0, 0.6% F: 20, 3.0, 20%</td>
<td>F: 13 (17), 33 (25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c: 18–75, 28.9 (10.4)</td>
<td>9 (116) 0.51 &lt; $d$ ≤ 0.69 all P &lt; 0.0001 c: 40% I: 2.3, 3.3, 1.2% M: 2.9, 5.3, 1.2%</td>
<td>All: 39% F: 0.1, 0.2, 0.2% F: 20, 3.0, 20%</td>
<td>F: 15 (19), 33 (25)</td>
<td></td>
</tr>
<tr>
<td>1000,² VBM</td>
<td>c: 18–74, 28.8 (14.3)</td>
<td>9 (116) 0.51 &lt; $d$ ≤ 0.69 all P &lt; 0.0001 c: 40% I: 2.3, 3.3, 1.2% M: 2.9, 5.3, 1.2%</td>
<td>All: 39% F: 0.1, 0.2, 0.2% F: 20, 3.0, 20%</td>
<td>F: 15 (19), 33 (25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c: 18–78, 26.8 (10.4)</td>
<td>9 (116) 0.51 &lt; $d$ ≤ 0.69 all P &lt; 0.0001 c: 40% I: 2.3, 3.3, 1.2% M: 2.9, 5.3, 1.2%</td>
<td>All: 39% F: 0.1, 0.2, 0.2% F: 20, 3.0, 20%</td>
<td>F: 15 (19), 33 (25)</td>
<td></td>
</tr>
<tr>
<td>1000,² VBM 18–26 subsample</td>
<td>c: 18–26, 21.5 (1.9)</td>
<td>9 (116) 0.46 &lt; $d$ ≤ 0.60 all P &lt; 0.0001 c: 55% I: 1.3, 0.8, 1.6% M: 0.8, 1.2, 0.5%</td>
<td>All: 53% F: 0.3, 0.4, 0.2% F: 10, 3.0, 10%</td>
<td>F: 16 (19), 33 (24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c: 18–26, 21.5 (1.9)</td>
<td>9 (116) 0.46 &lt; $d$ ≤ 0.60 all P &lt; 0.0001 c: 55% I: 1.3, 0.8, 1.6% M: 0.8, 1.2, 0.5%</td>
<td>All: 53% F: 0.3, 0.4, 0.2% F: 10, 3.0, 10%</td>
<td>F: 16 (19), 33 (24)</td>
<td></td>
</tr>
<tr>
<td>NKI, SBA, cortical thickness</td>
<td>c: 13–83, 41.0 (20.3)</td>
<td>7 (68) 0.41 &lt; $d$ ≤ 0.56 all P &lt; 0.0001 c: 26% I: 3.7, 3.6, 3.6% M: 3.7, 3.6, 3.6%</td>
<td>All: 24% F: 4.5, 2.0, 5.9% F: 20, 3.0, 20%</td>
<td>F: 21 (27), 33 (29)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c: 13–83, 41.0 (20.3)</td>
<td>7 (68) 0.41 &lt; $d$ ≤ 0.56 all P &lt; 0.0001 c: 26% I: 3.7, 3.6, 3.6% M: 3.7, 3.6, 3.6%</td>
<td>All: 24% F: 4.5, 2.0, 5.9% F: 20, 3.0, 20%</td>
<td>F: 21 (27), 33 (29)</td>
<td></td>
</tr>
<tr>
<td>NKI, SBA, volume</td>
<td>c: 12–85, 21.5 (2)</td>
<td>7 (68) 0.41 &lt; $d$ ≤ 0.56 all P &lt; 0.0001 c: 26% I: 3.7, 3.6, 3.6% M: 3.7, 3.6, 3.6%</td>
<td>All: 24% F: 4.5, 2.0, 5.9% F: 20, 3.0, 20%</td>
<td>F: 21 (27), 33 (29)</td>
<td></td>
</tr>
<tr>
<td>DTI fractional anisotropy</td>
<td>c: 17–43, 24.8 (4.6)</td>
<td>11 (116) 0.73 &lt; $d$ ≤ 1.05 all P &lt; 0.0001 c: 29% I: 2.9, 2.9, 2.9% M: 2.9, 2.9, 2.9%</td>
<td>All: 25% F: 2.2, 0.0, 4.3% F: 20, 3.0, 20%</td>
<td>F: 9 (15), 33 (33)</td>
<td></td>
</tr>
<tr>
<td>DTI connectivity</td>
<td>c: 17–43, 24.8 (4.6)</td>
<td>11 (116) 0.73 &lt; $d$ ≤ 1.05 all P &lt; 0.0001 c: 29% I: 2.9, 2.9, 2.9% M: 2.9, 2.9, 2.9%</td>
<td>All: 25% F: 2.2, 0.0, 4.3% F: 20, 3.0, 20%</td>
<td>F: 9 (15), 33 (33)</td>
<td></td>
</tr>
<tr>
<td>MADICS</td>
<td>c: 20–23, 21.6 (0.7)</td>
<td>7 (31) 0.43 &lt; $d$ ≤ 0.77 all P &lt; 0.0001 c: 64% I: 1.8, 1.1, 2.1% M: 0.8, 1.1, 2.1%</td>
<td>All: 59% F: 0.0, 0.0, 0.0% F: 20, 3.0, 20%</td>
<td>F: 17 (15), 33 (17)</td>
<td></td>
</tr>
<tr>
<td>ADD Health</td>
<td>c: 18–28, 22.2 (1.9)</td>
<td>8 (26) 0.41 &lt; $d$ ≤ 0.57 all P &lt; 0.0001 c: 81% I: 0.1, 0.2, 0.03% M: 0.0, 0.0, 0.0%</td>
<td>All: 70% F: 0.0, 0.0, 0.0% F: 20, 3.0, 20%</td>
<td>F: 27 (16), 45 (19)</td>
<td></td>
</tr>
<tr>
<td>Carothers &amp; Reis’ data</td>
<td>21.15 (7.68) 10 (10)</td>
<td>55% 1.0 &lt; $d$ ≤ 2.02 all P &lt; 0.0001 c: 48% I: 0.4, 0.9, 0.0% M: 0.4, 0.9, 0.0%</td>
<td>All: 55% F: 0.4, 0.0, 0.6% F: 20, 3.0, 20%</td>
<td>F: 11 (11), 48 (20)</td>
<td></td>
</tr>
</tbody>
</table>

$c$, males; $φ$, females.

*Statistically significant difference, $P < 0.0001$.

1If $000 = 1000$ Functional Connectomes Project.

If we were to include only regions with $|Cohen’s d| > 0.51$ as in the analysis of the entire sample, only three regions would have been included.

*Note the deviation from 33%, which reflects the inability to define the cutoff at 33% for some of the variables included in the analysis (see Methods for details).
Results
T1-weighted images of 169 females and 112 males (Table 1) were preprocessed and assessed for gray matter volume using voxel-based morphometry (VBM) (Fig. 1A and B). Of the 116 regions of gray matter defined using the Automated Anatomical Labeling atlas (AAL) (11) (Fig. 1C), 10 regions showing the largest sex/gender differences (|Cohen’s $d$| > 0.70, the largest |$d$| was 0.84; all $P < 0.0001$) were included in subsequent analyses (Table S1). Using the actual distributions of males and females in the sample, “male-end” and “female-end” zones were arbitrarily defined as the scores of the 33% most extreme males and females, respectively, and an “intermediate” zone was defined as the area in-between these two (Fig. 1D), we use the terms “male-end”/“female-end” as a shorthand for “the end of the continuum in which males/females are more prevalent,” respectively; note that our method of categorizing the continuum into three discrete classes inherently places some females at the “male-end” and some males at the “female-end”). Fig. 1E presents the gray matter volume of the 10 regions in each of the females (Left) and in each of the males (Right) using a color volume scale (as shown in Fig. 1D), and clearly illustrates the lack of internal consistency in most brains. The latter is also evident in Fig. 1F, which presents the number of “female-end” (x axis) and “male-end” (y axis) characteristics in females (red) and males (green). The circles at the (10,0), (0,10), and (0,0) coordinates represent individuals with only “female-end”, only “male-end”, or only “intermediate” characteristics, respectively; All other circles on the x and y axes represent individuals who have either “female-end” or “male-end” characteristics, as well as “intermediate” characteristics; The rest of the circles represent individuals with substantial variability, having both regions at the “male-end” and regions at the “female-end”. Thirty-five percent of brains showed substantial variability, and only 6% of brains were internally consistent (see Table 1 for more details). Notably, additional definitions of the “male-end” and “female-end” zones (50%, 20%, and 10%) similarly revealed a much higher prevalence of brains showing substantial variability compared with brains showing internal consistency (Table S2). Importantly, substantial variability is not a result of the overlap between females and males in each of the brain regions, as evident in Fig. S1A, which depicts the results that would have been obtained for these data under perfect internal consistency (for comparison, Fig. S1B–D depicts these data under internal consistency with different degrees of random noise, and Fig. S1E under no internal consistency). A similar pattern of results was obtained when we repeated the same analysis on the data of 495 females and 360 males obtained from the 1000 Functional Connectomes project (12) (Table 1, Table S2, and Fig. 2A). It is noteworthy that the two datasets revealed a similar pattern of results even though, of the 10 regions included in each analysis, only two were common (Table S1). To exclude the possibility that this pattern of results was a result of the large age range included in the two samples, as sex/gender differences have been reported to differ in different age groups (13–15), we repeated the same analysis on 625 individuals from the 1000 Functional Connectomes sample (385 females, 240 males) 18–26 y of age. Substantial variability was seen in 52% of brains, whereas internal consistency was evident in only 2.4% (Table 1, Tables S1 and S2, and Fig. 2B).

To test whether the pattern of results obtained in the two datasets was dependent on the type of analysis of the imaging data (VBM), and in particular on the “correction” for brain size included in the latter procedure, we performed the same analysis on a subgroup of the Nathan Kline Institute (NKI) enhanced sample (167 females, 100 males), whose T1-weighted images were preprocessed and manually corrected for cortical surface-based analysis. Using the FreeSurfer software package (Athinoula A. Martinos Center for Biomedical Imaging, Harvard University), 68 cortical regions were delineated in the native surface space (16), and the mean cortical thickness values (gray-white matter boundary to pial boundary) were calculated. This delineation method allows for direct comparison of regional and whole-brain cortical features in contrast to the VBM analysis described above, which includes normalization to a volumetric template space and “correction” for brain size. Also here, substantial variability was more common than internal consistency (24% and 10.5%, respectively; Table 1, Tables S2 and S3, and Fig. 2C). An analysis of the noncorrected volume of these 68 cortical regions, 23 subcortical regions of gray matter, and 77 regions of white matter revealed that substantial variability was much more common than internal consistency (23% and 2.2%, respectively; Table 1, Tables S2 and S3, and Fig. 2D). This latter finding is of particular interest because the sex/gender differences observed in this dataset were particularly large.

To test whether the pattern of results obtained was dependent on the type of imaging (T1-weighted images), we performed the same analysis on diffusion tensor imaging (DTI) indices of 69 females and 69 males. Using the AAL atlas used for the VBM, we calculated the average fractional anisotropy and mean diffusivity

Fig. 1. Assessing internal consistency in the human brain. (A) T1-weighted images were normalized and segmented using (B) the MNI template. (C) Voxels were mapped into 116 regions according to the AAL atlas. (D) The distribution of the gray matter volume in females (red) and males (green) of two of the region showing the largest sex/gender differences [left hippocampus (upper) and left caudate (lower)]. A continuous color presentation of the degree of “maleness” and “femaleness” was created separately for each of the 10 regions showing the largest sex/gender difference. Volumes falling in the “intermediate” zone are colored in white; volumes in the “male-end” and the “female-end” zones are colored using continuous blue-white and pink-white scales, respectively. (E) The degree of “maleness/femaleness” of each region for each female (Left) and male (Right). Each horizontal line represents the brain of one subject and each column represents a single brain region. The number above each column corresponds to the region’s number in the AAL atlas and in Table S1. (F) A bivariate scattergram (created using the Matplotlib library of Python) of the number of regions at the “female-end” (x axis) and at the “male-end” (y axis) in females (red) and males (green). The number of regions at the “intermediate” zone is not depicted because the number of “male-end”, “intermediate”, and “female-end” features always adds up to the number of features included in the analysis (which is the highest value on the x and y axes of the graph). The size of each circle is proportional to the percent of individuals from the same sex/gender category with an identical score on the two measures.
for each of the 116 regions of gray matter. The sex/gender differences in mean diffusivity were too small to survive the correction for multiple comparisons. Analysis of the fractional anisotropy data revealed substantial variability in 28% of brains and internal consistency in 5.8% (Table 1, Fig. 2).

We then used the DTI data to assess brain connectivity. Using deterministic fiber tractography, we estimated the connectivity strength between 90 regions of gray matter defined with the AAL atlas. Analysis of the 7 connections (of 4,005 connections) with the largest sex/gender differences revealed substantial variability in 48% of brains and internal consistency in only 0.7% (all of which were in the “intermediate” zone; Table 1, Tables S2 and S4, and Fig. 2F).

The low degree of internal consistency observed here in the human brain agrees well with studies demonstrating that humans often possess both “masculine” and “feminine” psychological characteristics (that is, personality traits, attitudes, interests, and behaviors that show sex/gender differences). Early attempts in the first half of the 20th century to measure masculinity-femininity using specially constructed scales have already revealed low or absent correlations between subscales measuring different characteristics of gender (17). Similar findings have led Janet Spence (18) to conclude that humans possess an array of masculine and feminine traits that cannot be captured using a uni-dimensional (masculinity-femininity) or a bidimensional (masculinity x femininity) model. However, to date, only a few studies have followed this line of research (19–21), and those that have, only assessed a small number of variables. We therefore used the same approach described above to analyze two open datasets that provide data on many psychological variables for a large number of subjects: the Maryland Adolescent Development In Context Study (MADICS) (22) and the National Longitudinal study of Adolescent Health (ADD Health) (23).

Of the different measures of behavior, personality characteristics, and attitudes available in MADICS, we analyzed data of seven variables with the largest sex/gender differences (Table S5) of 382 females and 188 males (Table 1). Substantial variability was evident in 59% of subjects and internal consistency in only 1.8% (all of which were in the “intermediate” zone; Fig. 2G and Fig. S2). Very similar results were obtained when analyzing the data of 2,239 males and 2,621 females from the ADD Health study, which is a study of a US-representative sample of adolescents. Substantial variability was evident in 70% of subjects and internal consistency in only 0.1% (all of which were in the “intermediate” zone; Fig. 2H, Table 1, and Table S5).

Last, we contrasted our analysis of the different brain- and gender-related datasets with a similar analysis of one of Carothers and Reis’ (22) behavioral datasets, which was unique in that it was the only behavioral dataset in which humans would be meaningfully grouped into two distinct categories on the basis of their sex (19). The dataset included 10 highly gender-stereotyped activities (e.g., playing videogames and watching talk shows; Table S5) specifically selected to differentiate between females (n = 157) and males (n = 106) of this subculture (introductory-level psychology class students at a large Midwestern American university). Accordingly, the sex/gender differences were very large (1.00 < |Cohen’s d| < 2.02), and the distribution of several of the variables was bimodal and highly skewed at both ends. This dataset was also unique in our analysis (compare Fig. 2I to Fig. 2A–H), in that there was almost no overlap between females and males in the possible combinations of the number of “female-end” and “male-end” characteristics, as in most subjects the number of “consistent” characteristics was higher than the number of “nonconsistent” characteristics. Interestingly, also in this sample, 55% of subjects showed substantial variability and only 1.2% were internally consistent (Table 1, Fig. 2I, and Fig. S2). In other words, even when considering highly stereotyped gender behaviors, there are very few individuals who are consistently at the “female-end” or at the “male-end”, but there are many individuals who have both “female-end” and “male-end” characteristics. Furthermore, although one’s sex is enough to predict whether this person would have more “female-end” or more “male-end” characteristics, it is not enough to predict this person’s specific combination of “female-end” and “male-end” characteristics (Fig. S2) (for further discussion of the question of prediction, see ref. 4).

**Discussion**

Consistent with previous findings (14, 15), our analysis of the structure of the human brain, which included most regions of gray and white matter, as well as measures of connectivity, revealed many nondimorphic group-level sex/gender differences in brain structure. There was extensive overlap of the distributions of females and males for all brain regions and connections assessed, irrespective of the type of sample, measure, or analysis (including analysis of absolute brain volumes). This extensive overlap underlines any attempt to distinguish between a “male” and a “female” form for specific brain features. Rather, the forms that are evident in most females are also the ones evident in most males (Fig. 1D). It is therefore more appropriate and informative to refer to measures of the brain in quantitative ways (Fig. 3) rather than in qualitative ways (e.g., “male”, “female” form). Another noteworthy observation is that the size of the sex/gender...
variable variability was not statistically different in any of the seven datasets, and the average proportion of “nonconsistent” characteristics was significantly higher in males compared with females in only one of the seven datasets (connectivity, P = 0.035). Thus, our findings that substantial variability is much more prevalent than internal consistency together with the lack of evidence for consistent sex/gender differences in the propensity to exhibit substantial variability do not support the classic view. Our findings are in line, however, with more recent thinking that masculinization and feminization are two independent processes and that sexual differentiation progresses independently in different brain tissues, “enabling genetically and environmentally induced variation in sexual differentiation of different tissues within a single brain” (4, p. 4; 10).

Our study demonstrates that although there are sex/gender differences in brain structure, brains do not fall into two classes, one typical of males and the other typical of females, nor are they aligned along a “male brain–female brain” continuum. Rather, even when considering only the small group of brain features that show the largest sex/gender differences, each brain is a unique mosaic of features, some of which may be more common in females compared with males, others may be more common in males compared with females, and still others may be common in both females and males. The heterogeneity of the human brain and the huge overlap between the forms that brains of males and brains of females can take can be fully appreciated when looking at the entire brain (Fig. 3 and Figs. S3 and S4).

In accordance with the brain data, our analyses of gender-related data revealed extensive overlap between females and males in personality traits, attitudes, interests, and behaviors. Moreover, we found that substantial variability of gender characteristics is highly prevalent, whereas internal consistency is extremely rare, even for highly gender-stereotyped activities (Carlson and Reis’ data). These findings are in line with previous reports that sex/gender differences in abilities and qualities are mostly nonexistent or small, that there is extensive overlap between the distribution of males and females also in behaviors, interests, occupation preferences, and attitudes that show larger sex/gender differences (24, 25), and that there are no or only weak correlations between gender characteristics (18, 20, 21). Thus, most humans possess a mosaic of personality traits, attitudes, interests, and behaviors, some more common in males compared with females, others more common in females compared with males, and still others common in both females and males.

Conclusions
The lack of internal consistency in human brain and gender characteristics undermines the dimorphic view of human brain and behavior and calls for a shift in our conceptualization of the relations between sex and the brain. Specifically, we should shift from thinking of brains as falling into two classes, one typical of males and the other typical of females, to appreciating the variability of the human brain mosaic. Scientifically, this paradigm shift entails replacing the currently dominant practice of looking for and listing sex/gender differences with analysis methods that take into account the huge variability in the human brain (rather than treat it as noise), as well as individual differences in the specific composition of the brain mosaic. At the social level, adopting a view that acknowledges human variability and diversity has important implications for social debates on long-standing issues such as the desirability of single-sex education and the meaning of sex/gender as a social category.

Methods
Data Collection and Preparation for Analysis.
Brain-related data. Data were obtained from four sources: Tel-Aviv University (the first brain-related dataset), University of Zurich (26) (DTI data), the 1000 Functional Connectomes Project (12), and the NKI enhanced sample

![Fig. 3. The human brain mosaic. The gray matter volume of all 116 regions of gray matter in females (Left) and in males (Right) from the first sample is represented using a continuous high-low (green-white-yellow) scale. Each horizontal line represents the brain of a single subject and each column represents a single brain region. The continuous high-low scale represents the relative volume of a brain region in a given brain relative to the volume of this brain region in all other brains (i.e., within a column). The regions that showed the largest sex/gender differences and were included in the internal consistency analysis are marked with a black bar. The number above each bar corresponds to the region’s number in the AAL atlas and in Table S1. (Inset) Magnification of a small part of a horizontal line (i.e., a single brain). The number in each colored cell is the volume of this region for this brain.](image-url)
Creating a Continuous Color Code.

Pink-white-blue. The scale was created separately for each brain region (and for each gender characteristic) based on the definitions of its “male-end”, “female-end”, and “intermediate” zones. Values in the “female-end” (“male-end”) zone were colored using the three-color scale conditional formatting function in Excel (Version 14.5.2), with the most extreme score defined as pink (blue), and the color bordering the “intermediate” zone defined as white (Fig. 1D).

Green-white-yellow. The scale was created separately for each brain region using the three-color scale conditional formatting function in Excel. The highest score was defined as green, the lowest score was defined as yellow, and the middle score was defined as white. In samples with equal numbers of males and females, the middle score was the median. In the other samples, the middle score was chosen so that the proportion of males on one side of this score equals the proportion of females on the other side to not bias the estimate of the middle of the distribution due to the nonequal number of males and females.

ACKNOWLEDGMENTS. We thank Dr. Bobbi Carothers (Washington University in St. Louis) and Prof. Harry Reis (University of Rochester) for allowing us to use their data and Prof. Reis for stimulating discussions of the mosaic hypothesis. This research used the Maryland Adolescent Development In Context Study of Adolescent Development in Multiple Contexts, 1991–1998 (Add Health dataset [made available in 2000, numeric data files]) that were collected by Jacqueline S. Eccles (Producer) and are available through the Henry A. Murray Research Archive of the Institute for Quantitative Social Science at Harvard University (Distributor). Special acknowledgment is due to Ronda R. Rindfuss and Barbara Entwisle for assistance in the original design. This research used data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullen Harris at the University of North Carolina at Chapel Hill and funded by Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant P01-HD31921, with cooperative funding from 23 other federal agencies and foundations. Information on how to obtain the Add Health data files is available on the National Longitudinal study of Adolescent Health (Add Health) website (www.cpc.unc.edu/addhealth). No direct support was received from Grant P01-HD31921 for this analysis. This work was supported by Swiss National Science Foundation Grants 320030-120661, 320030B-138668, 200303-138668, and 4-62341-05 and European Union Future and Emerging Technologies Integrated Project Presence: Research Encompassing Sensory Enhancement, Neuroscience, Cerebral-Computer Interfaces and Application (PRESENCIA) Grant 27731 to (L.J.).

5. Joel D (2012) Genetic-gonadal-gens-sex (3g-sex) and the misconception of brain and gender, or, why 3g-males and 3g-females have intersex brain and intersex gender. Biol Sex Differ 3(1):27.
7. Jordan-Young R, Rumiati RI (2012) Hardwired for sexism? Approaches to sex/gender characteristics. The Student t test was used to compare the mean proportion of “nonconsistent” characteristics in females and males, and the two-proportion z-test was used to compare the proportion of males and females showing substantial variability.