(Rethinking) the Neuroscience of Gender Identity

Simón(e) Sun, Ph.D.
Teddy G. Goetz, M.D., M.S.
Kristen Eckstrand, M.D., Ph.D.

Considering the role of psychiatry and mental health in the provision of mental health care for transgender, non-binary, and/or gender-expansive (TNG) individuals, it is natural to wonder what role the brain may play and what clues neuroscience may offer. For some, this interest may be driven by questions of “what makes me, me?”; others may wonder about how to harness the incredible power of the brain to address the mental health disparities plaguing TNG communities.

Despite its vast implications, the intersection between gender and neuroscience is only recently an area of inquiry, owing to the siloed histories of sex and gender in science, particularly across genetics, endocrinology, neuroscience, and psychiatry. Additionally, the concept of gender is tied to its origin in the psychiatrists and physicians who, when providing gender-affirming care, used gender to isolate psychological variation tied to sexual characteristics and maintain a cis-heteronormative binary understanding of sex. In this chapter, we describe the history of the neuroscience of sex and gender, current knowledge within neuroscience about sex and gender, and clinical applications of this knowledge.

Because the neuroscience literature has commonly (and incorrectly) conflated sex and gender (see later section, “The History of Gender in Neuro-
we introduce the following framework and approach, informed by numerous studies and commentaries (Figure 3–1) (Elliot and Richardson 2016; Fausto-Sterling 2019; Holmes and Monks 2019; Hyde et al. 2019; Llaveria Couselles 2021; Maney 2016; Miyagi et al. 2021).

1. **Sex** refers to biological components of an individual directly involved in sexual reproduction and the development thereof. Terms such as *female*, *intersex*, and *male* are used to describe characteristics of an individual—not individuals themselves. When referring to the sexual categorization of an individual, sex assigned at birth (male, female, intersex) is presumed.

2. **Gender/sex** refers to physiological and psychological aspects of an individual that are not directly related to but nonetheless are dependent on, interact with, or are influenced by sexual reproduction and its development. This aligns with a fieldwide shift toward precise language to distinguish when a phenomenon of study cannot be exclusively tied to either biological or sociocultural origin (Fausto-Sterling 2012; Kaiser et al. 2007; van Anders and Dunn 2009). Many studies that report sex differences in the brain are in this domain. Thus, we use gender/sex instead of sex when interpreting such findings. For example, a report on sex variability in a particular neuroanatomical region is referred to as a gender/sex difference. Terms such as *masculine*, *androgyne*, *ambiguous*, and *feminine* are used to describe gender/sex characteristics.

3. **Gender** refers to psychosocial and sociocultural aspects of an individual and their environment that relate to the perception by and of that individual as belonging to a sociocultural category with expected sex, gender/sex, and gender-related characteristics. As it is a unique aspect to human culture, the study of gender in animal models is inappropriate and inaccurate. Animal research that purports to study gender is relegated to category 2, gender/sex. Terms such as *non-binary*, *agender*, *man*, and *woman* are used to describe genders and their characteristics.

This framework recenters the question "what is the neuroscience of gender?" from unidimensional and static neuroanatomical correlations to the nonlinear interactions between biological, psychological, and sociological mechanisms at work (Llaveria Couselles 2021). It emphasizes the importance of the contextual dependence of sex labels (male, female, intersex) (Richardson 2022). In their scientific, medical, and cultural usage, these labels are often assigned based on one variable (e.g., in laboratory rodent studies, animals are often sexed by cursory visual estimation of anogenital distance and external genitalia). This is in contrast to the many biological contributors to sexual traits such as genetic and chromosomal makeup, hormone
fluctuations, developmental stage, age, environment, and social status. These contributors act as nodes of influence (bold words in Figure 3-1), in which specific mechanisms of study can reveal their influence on sex/gender-related characteristics.

Throughout this chapter, we acknowledge that many of the human studies we describe here were conducted with little regard for ethical considerations when involving TNG people as research participants. Although it is beyond the scope of this chapter to analyze and critique the failures of research programs centered on transness, it is necessary to approach these studies understanding that the research was conducted under the contexts emerging from normative cishepatriarchal white European colonialist gender norms. As a form of social structure, these neurosexist and racist data norms affect the perception of the researcher, which in turn affects data acquisition and interpretation (Duchesne and Kaiser Trujillo 2021; Schmitz and Höppner 2018). As we discuss this research, we aim to discuss these common pitfalls in neuroscience research yet also acknowledge the significant amount of work that remains to generate neuroscience literature that actually benefits TNG communities.

From a phenomenological standpoint, the history of sex/gender in science dates back as long as humans have been able to assign people to groups based on shared characteristics. The idea that there are fundamental differences between (binary) sexes that may be attributable to the brain is found in scientific writing from ancient Greece. Because the distinction between sex and gender was not evident in the scientific literature until the mid-1900s, studies were conducted without rigorous assessment of sex and assuming the existence of only two observable sexes. Such assumptions were evident in early neuroscientific research in the 19th century, which studied the brain based on visible observable differences in size, shape, and structure. This research falsely identified sex differences in brain size between assumed females and males, which were interpreted to reinforce the societally held belief that males were intellectually superior to females (Broca 1861). This research aligned with the neuropsychological ideas of sex prevailing at the time, according to which males who deviated from societal acceptable behaviors and emotions—particularly males who were not “masculine enough” or who engaged in same-sex relationships—were deemed “sexual inverts” (Ellis 1897). As the groundwork for modern neuroscience was not begun until the 20th century, critical scientific exploration about sex/gender in neuroscience did not occur until the 1950s.

(Rethinking) the Neuroscience of Gender Identity

The early 1900s brought advancements to the use of medicine to modify sex characteristics in what is now considered gender-affirming care, but the distinction between sex and gender identity was not described until the 1950s/60s (Stoller 1968). Researchers at the time hypothesized that sex and gender both existed on a binary, but that they were two nonoverlapping categories (Holmes and Monks 2019; Woolley 1910). Similarly, medical providers used the sex/gender distinction to support the emphasis on cis-heteronormative outcomes for intersex and transgender individuals seeking care (Gill-Peterson 2018). The result of this siloed biomedical discourse was an emphasis on sex as a product of nature and gender as a product of nurture or culture.

Unfortunately, the consequences of early definitions of sex as a biological variable and gender as a cultural influence affected which construct was deemed “scientific” and thus important to study in neuroscience. As a result, the majority of modern neuroscience has focused entirely on sex differences and sexual dimorphisms, albeit through flawed methodology that assessed neither sex nor gender clearly. Under such persistent neurosexist assumptions in research, neuroanatomists examined animal and human brains, concluding that there are macro- and micro-level sexual dimorphisms (Swaab and Hofman 1984), with the implications of these conclusions persisting today. For example, based on early (unreplicated) findings of differences in brain hemispheric laterality, popular culture pundits mused that men are more “right-brained” and women more “left-brained” to explain differences in emotionality. Despite recent meta-analyses demonstrating that assumed sex accounts for <1% of the variation in the brain (Eilott et al. 2021), one need only look to contemporary culture to find hundreds of books and articles falsely explaining, and reinforcing, binary social gender assumptions through poorly sourced neuroanatomical studies.

Although much of the fanfare over sexual dimorphisms is unwarranted, there are notable sexual dimorphisms in the brain, and beginning in the 1980s, these regions became the key focus for identifying the etiology of TNG identities. In contrast to emerging social and philosophical theories of the time suggesting that gender is not binary (Butler 1990), neuroscience applied century-old hypotheses of sexual inversion to develop the “cross-sex shift hypothesis” of the brain. Under this antiquated hypothesis, brains exposed to various factors (e.g., genetics, prenatal hormones) could develop to become more “male-like” or “female-like,” and these cross-sex shifts were supposed to explain why someone assigned one sex at birth may have an incongruent gender (Swaab et al. 2021) or even exhibit same-sex sexual behavior (Abe et al. 2021).

After nearly 30 years of neuroscience research, such ideas persist today in the scientific literature across domains of sexual orientation, gender identity,
and gender expression (Abé et al. 2021; Folkerska–Zawowska et al. 2020; Garcia-Sifuentes and Maney 2021) despite small samples, small effect sizes, failure to replicate effects in larger samples (Ganna et al. 2019; Lambert 2019), and repeated calls for methodological improvements and reconsiderations (Clayton and Tannenbaum 2016; Miyagi et al. 2021; Woitowich and Woodruff 2019). Essentialist assumptions regarding the etiology of gender identity have led to overinterpretation of results, including attempts to use research to describe negative health trends in lesbian, gay, bisexual, transgender, queer, intersex, asexual, and more (LGBTQIA+) populations (Li et al. 2021). These conclusions are incompatible with data demonstrating that inadequate social support and gender affirmation results in significant improvements in the mental health of TNG people (Colizzi et al. 2013; Durwood et al. 2017; Nobili et al. 2018).

The history of gender in neuroscience has not yet led to appreciable benefits for TNG communities, but there is now a robust evidence base debunking a singular neural theory for sex, gender expression, or gender identity. Such evidence continues to support sex and gender as having complex, multifactorial, individual characteristics that are not amenable to simple explanation. Further, based on the lack of meaningful findings from the last 150 years, one could argue that the primary lesson we have learned from neuroscience is that continued application of binary assumptions of sex and gender in neuroscience will be exceedingly unlikely to lead to scientific advancement for people of any gender, much less those who are TNG (Aghì et al. 2022).

A NEUROSCIENCE OF GENDER: INTEGRATING NEUROENDOCRINOLOGY

With these considerations in mind, the overlap of different copulations relating to gender and sex in transgender and intersex populations suggests that two particular biological and physiological factors—genes and hormones—significantly contribute to an individual’s sense of gender.

The prevalence of gender dysphoria in intersex individuals is on average higher than in the general population. Importantly, intersex individuals with 17β-hydroxysteroid dehydrogenase deficiency and 5α-reductase deficiency have among the highest incidence of gender dysphoria (Babu and Shah 2021; Furtado et al. 2012). Conversely, transgender persons who have sought medical treatments for gender dysphoria are more likely to have atypical karyotypes (chromosomal profiles), with differences presenting in both sex chromosomes and autosomes (Inouhi et al. 2011). Recent studies have found pre-

Luminary evidence that transgender individuals have genetic variants in genes related to hormone signaling, including hormone receptors, steroidogenesis enzymes, and related cofactors (Fernández et al. 2020; Foreman et al. 2019; Theisen et al. 2019); however, the replicability of these data remains to be seen. Nevertheless, these preliminary findings, in conjunction with the success of gender-affirming hormone therapy (GAHT) for TNG populations (Aghì et al. 2022), suggest that intrinsic physiological aspects of gender/sex (Figure 3–1, left section) emerge in part from the interaction of steroid hormones, gene expression, and neural circuitry in the brain.

A large proportion of gender/sex characteristics in the brain are established and regulated by the “gonadal” or “sex” hormones: the estrogens (estrone, estradiol, estril), progesterone, and testosterone. All derived from cholesterol, steroid hormones circulate in the bloodstream and can freely cross the blood–brain barrier. Primarily produced by gonads, adrenal cortex, and placenta, they are also synthesized locally in the brain (Do Rego et al. 2009). The extent to which neurally synthesized hormones contribute to neurological development, differentiation, and function is still underexplored. Several populations of neurons also express aromatase and locally convert circulating and neuron-synthesized testosterone into estrogen to mediate their effects (Naforlin and Ryan 1975). Some prehormones are found at much higher levels in the brain than in the plasma (Corpechot et al. 1981; Eber et al. 2006; Labrie et al. 2005), and the steroidogenic enzymes or their mRNAs have been found in both neurons and glial cells in various brain regions (Do Rego et al. 2009).

Steroidal hormones are ligands that readily pass through cellular and nuclear membranes. Each hormone has a set of receptors that it can bind to, each encoded by its own gene: estrogen bind to estrogen receptor alpha (ERα, encoded by ESR1) and/or beta (ERβ, encoded by ESR2), progesterone to progesterone receptor (PR, encoded by PGR), and testosterone to androgen receptor (AR, encoded by Ar). Thus, the expression of these nuclear receptors determines whether a brain region can respond to the corresponding hormone.

When a hormone–receptor complex is formed, it acts as a transcription factor to regulate the expression of target genes. The hormone–receptor complex interacts with other nuclear receptor coregulators (Sun and Xu 2020) and transcriptional machinery that bind directly to DNA at specific nucleotide sequences called response elements: estrogen-ERα/ERβ at estrogen response element (ERE) and testosteron-AR at androgen response element (ARE). Testosterone can also bind to glucocorticoid response elements (GRs), which typically are bound by corticosteroids. Glucocorticoid receptors can also bind to ARs (Manoli and Tollekohu 2018). These receptors can also tether to other DNA-binding proteins and can regulate gene expression
without interacting directly with the genome. Hormone receptor action through DNA binding is known as canonical hormone signaling. Noncanonical hormone signaling, the effects of which are only partially understood, is known to occur through pathways that do not involve receptors interacting with DNA or other transcription factors (Davey and Grossmann 2016; Heldring et al. 2007). Because they have similar structures, progesterone is also able to bind to androgen receptors (Bardin et al. 1983) and glucocorticoid receptors (Bupprecht et al. 1993); dihydrotestosterone (DHT) is an even more potent AR agonist than testosterone (Tóth and Zakár 1982).

In addition to binding to DNA, hormone receptors can regulate the epigenome to control gene expression. Epigenetics refers to the control of gene expression by molecular changes to DNA and/or its 3D structure, without changing the DNA sequence itself. There are two primary ways in which epigenetic changes can occur: DNA methylation and chromatin remodeling.

DNA methylation involves the addition of a methyl group onto a cytosine nucleotide that is upstream of a guanine to activate or repress gene expression. In the brain, DNA methylation is mediated by DNA methyltransferases.

Epigenetic control also occurs through chromatin remodeling. Chromatin is a mixture of DNA and proteins that form chromosomes. DNA wraps around proteins known as histones, forming units of chromatin known as nucleosomes. The accessibility of chromatin controls whether a gene can be expressed. Heterochromatin, tightly packed chromatin, prevents gene expression machinery from accessing genes. In contrast, euchromatin is open and loosely packed, exposing DNA sites for gene expression machinery to bind to and initiate gene transcription. Control of chromatin accessibility occurs through posttranslational modifications (PTMs) to histone proteins, which affect the electrostatic interactions between DNA and histones. PTMs such as acetylation, methylation, and phosphorylation dynamically change in response to neural activity.

It is hypothesized that sex differentiation of neural tissue is mediated by permanent epigenomic effects of perinatal hormonal surges. Indeed, several studies have identified genes that are differentially expressed by sex (Kurian et al. 2010; Nugent and McCarthy 2011; Westberry et al. 2010), as well as genomewide sex variability (Ghaemmaghami et al. 2014; Nugent et al. 2015). These results, along with recent work identifying neuronal gene regulation by ERα (Gengenbacher et al. 2022), illustrate the dynamic nature of epigenetic regulation by hormones, suggesting that some aspects of early life organization are not completely permanent. The functional consequences of hormonal regulation of gene expression, as well as hormonal alterations to the activity of neural networks, remain poorly understood.

Contemporary understanding of the pattern and development of gonadal hormone receptor expression in the mammalian central nervous system is incomplete. Postmortem human tissue studies have revealed the presence of hormone receptors in the brain; studies in animals, particularly rodents, have provided a map of neuronal hormone receptor expression at a relatively high spatiotemporal resolution. The four canonical gonadal hormone receptors—ERα, ERβ, PR, and AR—are highly expressed in the limbic system—medial amygdala, bed nucleus of the stria terminalis, medial preoptic area of the hypothalamus, and ventromedial hypothalamus—irrespective of gonadal or chromosomal sex, although there is notable sex variability. These brain areas are core components of numerous neural circuits that mediate reproductive, territorial, aggressive, parenting, and social behaviors, commonly referred to as the social behavior network (Li and Dulac 2018; Newman 1999; Petrisul et al. 2017; Wei et al. 2021). The broadly successful outcomes of combined social and medical gender affirmation suggest that these regions mediate important components of gender identity.

ERα, AR, and PR mediate feeding and energy homeostasis (Mauvais-Jarvis et al. 2013), and are found in the arcuate nucleus as well as in the midbrain regions such as the periaqueductal gray, raphe nucleus, substantia nigra, and ventral tegmental areas (Creutz and Kritzer 2004; Mitra et al. 2003; Purves-Tyson et al. 2012; Quadros et al. 2008; Quesada et al. 2007; Shughrue et al. 1997). These receptors are also found in the neocortex (important for sensory processing, motor execution, and cognitive functions) and the hippocampus (involved in learning and memory) (Cara et al. 2021; Mitra et al. 2003; Núñez et al. 2003; Sar et al. 1999; Shughrue et al. 1997). Individual neurons can express varying combinations of receptors.

Hormone receptor expression is known to change with age, which further complicates understanding the implications of spatiotemporal distribution. The mechanisms by which gonadal hormones act on the brain have been explored primarily in the context of how long-lasting neurobehavioral sex differences develop. The earliest of these experiments, reported by Phoenix and colleagues in 1959, demonstrated that testosterone administration to pregnant guineas pigs resulted in a reduction in female-typical sexual behavior (lordosis) and an increase in male-typical sexual behavior (mounting) in female offspring (Phoenix et al. 1959). Based on their results, they proposed the Organization-Activation Model (OAM) of hormone-mediated neurobehavioral sex differentiation. The OAM posits that hormones sexually differentiate behaviors by permanently organizing brain tissue during early critical periods of development; in adulthood, these hormones play an activating role on the differentiated brain, resulting in sex variability of behavior. Since then, similar results have been reported in mice, rats, nonhuman primates, and humans, mediated by prenatal or perinatal exposure to androgens (or in the case of rodents, estrogens by way of aromatization of testosterone) (Hines 2011; McCarthy et al. 2009).
Gender-Affirming Psychiatric Care

This research is broadly related to gender, but the attempts to translate such theories for TNG and intersex health care (Diamond 2009; Guillaume et al. 2016; Shirazi et al. 2020; Zucker et al. 1996) have been detrimental and exploitative to the very populations purported to benefit from the work (Colapinto 2006; Gill-Peterson 2018). Additionally, the OAM cannot capture changes outside development and the ontological necessity of gonadal hormones to maintain sexual differentiation (Gegenbauer et al. 2022; Knodler et al. 2022), which makes necessary a more refined dynamic model of neuroendocrine function to untangle the neurological mechanisms related to gender.

Conclusion: What does this mean for psychiatry in practice?

Even in a test tube, gender is messy! Considering real patients, in all of their life experiences, only underscores this messiness. Thus, in the final portion of the chapter, we ask you to think about gender with intentionality, as a multidimensional factor for each of your patients.

- **Hormonal milieu**: How does your patient's dominant hormonal milieu influence their evidence-informed psychiatric care? How would initiating or ceasing GAHT change treatment?
- **Societal illegibility and associated stress**: When and how does your patient experience minority stress related to their gender? Expressing their gender presentation at home, with roommates, with family, at work, walking down the street? Seeking a legal name or gender marker change? Seeking medical care? How is their societal illegibility changing over time? How can these answers guide psychotherapy interventions?
- **Gender dysphoria and its impact on psychiatric diagnoses**: If applicable, how is your patient's gender dysphoria affecting their mental health? How can you work within an interdisciplinary treatment paradigm to best support gender-affirming care, in conjunction with psychiatric interventions? (See Chapters 4, 15, and 21 for information on psychopharmacology, eating disorders, and surgery evaluations.)
- **History of trauma and microtraumas**: What kinds of traumas and microtraumas has your patient experienced and do they currently experience? How do these stressors affect their psychiatric illness? How can this knowledge, along with hormonal milieu, inform their treatment?

All of these experiences affect the brain on multiple levels—genetics, neural connectivity, neural activity—and all are intrinsically rooted in the way that a patient experiences their gender within society. This means that when we are making psychopharmacological decisions for patients, we should be thinking about how to best work responsively with regard to all layers of a patient's gender, including their gender-affirming health care and the ways in which minority stress might be impeding the molecular actions of the drugs we prescribe (see Chapters 2 and 11 for information on minority stress and trauma). During psychotherapy with a patient, we must keep in mind both how their hormonal milieu may influence their neurotransmitter functioning and how their anxiety about their gendered legibility might be distracting them and preventing them from engaging fully with the therapy process, or from forming memories and making connections as you would typically expect from such an interaction.

Most of these clinical considerations, while intuitive, remain hypothetical in the sense that historically, research time, funding, and resources have not been devoted to exploring how to provide equitable mental health care for all genders. We imagine the potential power of future clinical trials focused on developing drugs that work synergistically with a specific hormonal milieu, designed for and tested in TNG patients. We hope that the discussion in this chapter inspires future work to advance equity and justice in pharmacotherapy.

Recognizing this profound opportunity for future research—rethinking the neuroscience of gender identity and how to apply it to psychiatric care for TNG patients—we propose a few particular areas in which we hope to see the research literature expand over the next few years (for more specific discussion, please see Aghi et al. 2022):

- **Impact of hormonal milieu on affect and behavior**: Because estrogens, androgens, and progestogens cross the blood–brain barrier and have receptor targets present in the central nervous system, absolute and relative hormonal levels likely affect mood, arousal, and responses to stimuli. This area of research would be fruitful for better predicting affective and behavioral changes associated with GAHT, beyond the psychological impact of greater congruence between self-understanding and embodiment (Hughto et al. 2020; Murad et al. 2010). These areas should also be studied for individuals who undergo pubertal blockade to halt endogenous puberty before initiating GAHT.
- **Possibilities of modulating GAHT outcomes**: Gender-affirming care is far from one-size-fits-all (e.g., Beek et al. 2015), and GAHT is capable of driving a variety of embodied changes, some desired by an individual and others adamantly unwanted. Methods for modulating specific hormonal effects through more selective hormone therapy and/or neuro-modulation mechanisms have the potential to revolutionize GAHT and
allow TNG people to more directly pursue their affirmation goals and seek gender euphoria.

- **Impact of dominant hormonal milieu on psychiatric medication efficacy**: In light of the receptor distribution for estrogens, androgens, and progestogens in the brain, it is possible (perhaps likely) that psychiatric medication efficacy may be modulated by the hormonal environment. Some medications may be more likely to achieve the clinically desired results for certain individuals, based on their hormonal milieu, and the success of these medications may change over time. Research on differential responses to psychiatric medications between groups stratified by dominant hormonal milieu may offer improved guidelines for first-line medication selection on an individual basis.

- **Impact of chronic stress from gender-related micro- and macroaggressions on neurochemistry, psychiatric illness, and response to mental health treatment**: Physiological and psychological impacts of chronic stress are well characterized (Agorastos and Chrousos 2022), and TNG people are subject to elevated levels of chronic stress because of marginalization related to gender identity and gender modality (see Chapters 2 and 11 for information on minority stress and trauma). Accordingly, experiences of discrimination and harassment may affect neurochemistry and mental health, as well as how an individual responds to a given treatment. Such areas warrant further investigation.

- **Impact of gender-affirming surgery on neurochemistry, psychiatric illness, and response to mental health treatment**: As discussed earlier, exposure to chronic stresses is known to increase risk of psychiatric illness; a chronic sense of incongruent embodiment could qualify as such a stimulus. Indeed, improvements in mood, anxiety, and suicidality have been demonstrated following gender-affirming surgery (e.g., Akhavan et al. 2021; Hughot et al. 2020). This effect may be mediated by induced alterations in neurochemistry, which could even change how an individual responds to specific medications. These areas have not yet been explored.

**REFERENCES**


Clayton JA, Tannenbaum C: Reporting sex, gender, or both in clinical research? JAMA 316(18):1863–1866, 2016 27802482


(Rethinking) the Neuroscience of Gender Identity
Llaveria Caselles E: Epistemic injustice in brain studies of (trans)gender identity. Front Sociol 6:608328, 2021 33869551

(Rethinking) the Neuroscience of Gender Identity

Gender-Affirming Psychiatric Care

Fausto-Sterling A: Gender/sex, sexual orientation, and identity are in the body: how did they get there? J Sex Res 56(4):529–555, 2019 30875248
Gill-Peterson J: Histories of the Transgender Child. Minneapolis, MN, University of Minnesota Press, 2018
Gender-Affirming Psychiatric Care

Richardson SS: Sex contextualism, Philos Theory Pract Biol 14(2), 2022
Stoller RJ: Sex and Gender: The Development of Masculinity and Femininity. Abingdon, UK: Routledge, 1968
Sun Z, Xu X: Nuclear receptor coactivators (NCOAs) and corepressors (NCOBs) in the brain. Endocrinology 161(4):bpa0083, 2020 32449767

(Rethinking) the Neuroscience of Gender Identity

Woolley HT: A review of the recent literature on the psychology of sex. Psychol Bull 7(10):338–342, 1910