

CONSIDERATIONS FOR THE DEVELOPMENT AND IMPLEMENTATION OF  
TRANSGENDER – INCLUSIVE GENDER DEMOGRAPHIC QUESTIONS

Margaret L. Haupt

Submitted to the faculty of the University Graduate School  
in partial fulfillment of the requirements  
for the degree  
Doctor of Philosophy  
in the Department of Psychological and Brain Sciences,  
Indiana University  
April 2019

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Doctoral Committee

---

Eliot R. Smith, PhD

---

Mary C. Murphy, PhD

---

Julia R. Heiman, PhD

---

Justin R. Garcia, PhD

March 26, 2019

## Acknowledgements

Thank you to the Society for the Psychological Study of Social Issues for the Clara Mayo Grant, which funded participant payments for Study 1.

I am extremely grateful to my advisor, Dr. Eliot R. Smith, for allowing me the scholarly freedom to conduct this project in the first place. Since my very first phone interview with him, he has been quietly supportive and unassumingly kind. I am so grateful for his help in keeping my project scope manageable and the big picture clear, his assistance with statistical analyses, and his ability to return thoughtfully annotated chapter drafts in a matter of days. He is the most diplomatic, constructive critic I have ever had, and I always leave his office feeling better than when I went in.

I am also deeply thankful for my committee members. Dr. Mary Murphy helped me sketch out the first designs of the study that (five years later) became Chapter 4, and taught me so much about the complex nature of prejudice. Dr. Justin Garcia was a ray of sunshine and positive energy in every class and committee meeting, kept me anchored in the world of sex science and the Kinsey, and encouraged collaborations that led to two publications. Dr. Julia Heiman helped to shape the ideas that eventually become Chapter 7, asked me hard questions, and was an invaluable voice reminding me of the practical constraints facing clinicians.

Several special people helped me feel so much more at home in academia, and without their encouragement I would never have finished this dissertation. I am so grateful to Dr. Anna Pope, for understanding and believing in me and this project within five minutes of meeting me, to Dr. Amanda Gesselman for no-nonsense advice and airport camaraderie, to Stephanie Lochbihler for going through the same things at the same times with the best of humor, and to Dr. Lisa Thomassen, for realism, good humor, constant check-ins, and positive reinforcement.

Equally essential were my partners, Zim and Julia, who provided constant love and support. I am so thankful that they were willing to incessantly discuss survey design and the meaning of gender identity for the past six years, and that they agreed to spend that time in Indiana with me.

I am deeply indebted to my parents, Cathy and Dan, for their constant love, and for providing me with such strong early role models of compassion, critical thinking, and social conscience. They have always encouraged me to test my intellectual limits and choose my own life path, and believed in my ability to rise to meet any challenges I might face.

There are so many other people who helped me with the technical and theoretical aspects of this project, and provided support and encouragement throughout this process. Dr. Erick Janssen helped to get me involved at the Kinsey, and Dr. Stephanie Sanders' *Gender, Medicine, and the Body* class helped to clarify my thinking on much of the gender theory in this paper. My friend Kristin Brethova helped me schedule invaluable dissertation work days leading up to the end of the project. My research assistants Tasha Smith, Ell Thomas, Gwendolyn Evans, Eric Budzielek, Jess Mao, Michael Spors, and Madison Kaufman helped me to talk through this project, pilot test early survey drafts, and code comments. Countless conversations with students, faculty, and conference attendees also contributed to this final product.

Thank you all!

Margaret L. Haupt

CONSIDERATIONS FOR THE DEVELOPMENT AND IMPLEMENTATION OF  
TRANSGENDER – INCLUSIVE GENDER DEMOGRAPHIC QUESTIONS

Binary gender questions are often criticized for their inability to identify gender minorities (i.e. transgender and non-binary people) and their inconsistency with modern multidimensional theories of gender. Several inclusive gender questions have been proposed to address these problems, but the evidence supporting their use is less conclusive than many organizations have claimed, and the existing body of literature is generally lacking in theoretical grounding. This dissertation aims to partially address this gap between theory, measure development, and measure application.

Chapter 1 explores previous conceptualizations of gender from a variety of disciplines and explains why the binary gender question is theoretically and pragmatically inadequate for assessing gender and transgender status, while Chapter 2 reviews existing literature proposing alternatives to the binary gender question. Chapters 3 through 5 provide novel empirical evidence from two studies about the *simplicity*, *acceptability*, and *validity* of the most-commonly recommended gender identity measures. Chapter 6 shifts focus from transgender participants to society at large, testing the hypothesis that some gender question formats may cue gender essentialism, and that revising them might thus reduce prejudice toward transgender people. Finally, Chapter 7 describes an attempt to extend transgender-inclusive questions to the context of sexual orientation, a construct which has historically been grounded in a binary view of gender.

---

Eliot R. Smith, PhD

---

Mary C. Murphy, PhD

---

Julia R. Heiman, PhD

---

Justin R. Garcia, PhD

## Table of Contents

Introduction and Overview .....	1
Chapter 1: What’s Wrong With the Gender Binary? .....	4
Chapter 2: Alternatives to the Binary Gender Question .....	25
Chapter 3: Simplicity of Inclusive Gender Identity and Transgender Status Measures .....	53
Chapter 4: Acceptability and Identity Safety of Gender Measures .....	113
Chapter 5: Validating Inclusive Gender Identity Measures .....	161
Summary of Chapters 3-5 .....	208
Chapter 6: Effects of Gender Identity Measures on Attitudes toward Gender Minorities .....	215
Chapter 7: Including the Gender Spectrum in Measures of Sexual Orientation.....	230
References .....	273
Appendix A .....	293
Appendix B .....	299
Appendix C .....	320
Appendix D .....	372
Margaret Haupt CV	

## Introduction and Overview

Gender<sup>1</sup> is perhaps the most ubiquitously collected demographic information, requested by government officials, physicians, academics, and laypeople alike. Despite the vast array of situations in which gender information is requested and the incredibly diverse applications for which this information is used, most gender questions look remarkably alike. Indeed, many survey designers consider the traditional gender question “Male or Female?” so familiar and self-explanatory that the options are commonly abbreviated M/F to save space. In recent years, however, this ubiquitous “Male or Female?” gender question (hereafter referred to as the *binary gender question*) has come under scrutiny from a variety of independent scholars and professional and governmental organizations. Recognizing the theoretical gap between current understandings of gender and the ubiquitous use of the binary gender question, the American Psychological Association (2009), the Institute of Medicine (2011), the Department of Health and Human Services (2013) have each called for human subjects research to routinely include a standard set of inclusive gender demographic questions which can identify gender minorities (i.e. binary and non-binary transgender people as well as gender non-conforming people).<sup>2</sup> However, none of these organizations have put forward candidates for such an inclusive gender measure, instead calling on researchers to develop one.

In response, transgender-advocacy organizations such as the Center for Excellence in Transgender Health and the Williams Institute, as well as individual researchers, have proposed a variety of replacements for the binary gender question (Sausa et al., 2009; Tate et al., 2013;

---

<sup>1</sup> The term “gender” can refer to any of a wide range of constructs, and there is no scientific consensus on a definition for the term (Muelenhard & Peterson, 2011). Throughout this paper, I use the term “gender” to refer in a general sense to the fluctuating collection of constructs which are currently described as “gendered” in academic literature and popular culture (see Hegarty & Pratto, 2010 for rationale).

<sup>2</sup> The political climate has changed substantially since this project began.



GenIUSS Group, 2014; Cahill & Makadon, 2014; Reisner et al., 2014). While virtually all of these methods would be improvements over the status quo, the evidence supporting any particular measure is less conclusive than many organizations have claimed, and the existing body of literature is generally lacking in theoretical grounding.

In this dissertation, I begin by describing the reasons why the binary gender question is theoretically and pragmatically inadequate for assessing gender identity and transgender status, and providing an overview of existing models of these constructs (Chapter 1). I then identify the most promising alternative measures which have already been put forward, and the evidence for and against their use based on previous research (Chapter 2).

The second part of the dissertation (Chapters 3 through 5) explores the *simplicity*, *acceptability*, and *validity* of the most-commonly recommended gender identity measures, presenting results from different parts of the same two studies (each sampling undergraduates and self-identified gender minorities for four total samples). Chapter 3 considers the *simplicity* of these questions. I descriptively compare participants' responses to the most commonly recommended trans-inclusive gender identity questions, and offer practical recommendations to researchers for simplifying and analyzing the often-complex data yielded by these inclusive questions. Chapter 4 focuses on *acceptability*, testing the hypothesis that participants (especially those who are transgender and/or non-binary) experience identity threat when seeing certain types of gender questions, that this identity threat may reduce their liking of researchers and willingness to participate in research, and that the threat posed by a given question may differ by context. Chapter 5 concerns *validity*. I present evidence of the substantive validity, generalizability, and external validity of the multiple-choice gender identity measure (alongside

the transgender identity measure), as Chapters 3 and 4 suggest that it offers the best combination of simplicity and acceptability.

The final section of the dissertation explores two directions for future research. Chapter 6 shifts focus from transgender and non-binary participants to society at large, testing the hypothesis that some gender question formats may cue *gender essentialism*, a belief that of members of a gender category share an underlying natural, innate “essence” (e.g. Haslam, Rothschild, & Ernst, 2000). Gender essentialism, especially a belief in a biological gender binary, may lead to increased transprejudice and negative stereotyping of transgender and gender non-conforming people (Ching & Xu, 2017). Finally, Chapter 7 describes an attempt to extend transgender-inclusive questions to the context of sexual orientation, a construct which has historically been grounded in a binary view of gender. I present results from a series of studies testing a basic trans-inclusive sexual orientation measure, and suggest directions for future study.

Overall, I argue that a successful gender demographic question is one that balances *simplicity, acceptability, and validity*. That is, a gender demographic question should be easy for researchers to use in data collection and analysis, accurately predict the construct(s) of interest, and be understandable and inoffensive (perhaps even welcoming) to diverse participants. While no question can perfectly satisfy all these requirements, some options are better than others, and I put forward a few promising candidates in the hopes of increasing the general quality of gender demographic questions used in social science research.

## Chapter 1: What's Wrong With the Gender Binary?

Western culture has traditionally perceived gender as binary, composed of two and only two mutually exclusive categories: “male” and “female.” Both laypeople and academics frequently subscribe to biological determinism when it comes to gendered psychological and social constructs. Perhaps the clearest example of this sort of thinking occurs after the birth of a child, when the declaration “It’s a boy!” or “It’s a girl!” follows a physician’s cursory inspection of a child’s genitals (Hubbard, 1996). In our cultural imagination, a person’s genitals (but not their gonads, hormones, or chromosomes) are viewed as the essential determinant of their gender (e.g. Kessler & McKenna, 1978, Bettcher, 2007). Indeed, the belief that genitals are the determining sign of gender is so deeply rooted that physicians have justified cosmetic genital surgeries on intersex infants with claims that the children will otherwise grow up without a clear gender identity<sup>3</sup> (e.g. Kessler, 1990; Beh & Diamond, 2000; Blackless et al. 2000). For all its complexity, the vast majority of social science research still reflects this unspoken assumption that children with penises always grow up to be men, while children with vaginas always grow up to be women.

---

<sup>3</sup> This practice was originated by John Money, the same researcher who created the distinction between sex and gender. While he believed that gender identity was entirely socially determined, he argued that “a child should look like the sex in which he or she is assigned. The child and the parents, to say nothing of relatives and friends, need to be able to see that the genitalia do not tell a lie” (Money 1974, p. 216). According to this line of reasoning, genitals determine gender identity because the lay public is certain that genitals are the true source of gender identity. However, Money also placed great emphasis on a child’s ability to engage in normative heterosexual sex upon reaching adulthood, arguing that “in the final analysis the morphology of the external genitals and their surgical and endocrine modifiability are the most important criteria. It is useless to condemn a child to grow up in, and to differentiate the gender identity of, the sex in which he can never function coitally” (Money 1974, p. 221). This assumption of a linkage between gender identity, sexual orientation, and sexual practice, and the concomitant emphasis on both heterosexuality and the suitability of genitals for normative heterosexual intercourse continues to pervade research and theorizing on the care of intersex and transgender people (e.g. Kessler, 1990; Dreger, 1998; Lawrence, 2010; Cohen-Kettenis & Pfafflin, 2010; Drescher, 2010). Many transgender and intersex people have reacted strongly against this theoretical approach, and are wary of questions about sexual orientation as a result (e.g. Cohen-Kettenis & Pfafflin, 2010; Serano, 2010). While a full consideration of these issues is outside the scope of this paper, it is important to be aware of this tension and distrust when deciding what demographic questions to include in studies and how to frame them to avoid inadvertently alienating participants.

Despite the strong cultural tendency to equate them, the assertion that gender identity and genitals are theoretically distinct constructs has appeared in various forms for nearly a century in scholarly literature from medicine (e.g. Hirschfeld, 1923; Money et al., 1972), sociology (e.g. Garfinkel, 1967; Hines & Sanger, 2010), gender studies (e.g. Butler, 1990; Serano, 2007; Fausto-Sterling, 2012), developmental psychology (e.g. Egan & Perry, 2001; Diamond et al., 2011), clinical psychology (e.g. Singh et al., 2010; Zucker, 2012; Joel et al., 2014), and social psychology (e.g. Olson et al., 2015; Tate et al., 2014) among other disciplines, although each field has approached this idea somewhat differently. Perhaps most influential was the distinction between sex (biological factors) and gender (social factors) introduced by John Money and colleagues in 1955. By 1972, they described chromosomal, gonadal, and hormonal sexes as well as gender identities and roles (Muehlenhard & Peterson, 2011). Subsequent research has made clear that chromosomes, gonads, hormones, genitals, and secondary sex characteristics can each take multiple forms (e.g. Blackless, 2000; Fausto-Sterling, 2000; Consortium on the Management of Disorders of Sex Development, 2006). While certain combinations are more common, the form of any one of these sexed physical characteristics cannot perfectly predict the forms of the others.

Gender identity is no less complex. While binary gender questions presume (and thus find “evidence” for) the existence of two and only two genders, other types of questions find evidence for multiple genders (e.g. Factor & Rothblum, 2008; Johnson & Wassersug, 2010; Kuper et al., 2012) and/or a gender spectrum (e.g. Joel et al., 2014). However, the precise constructs signified by gender remain ambiguously defined (and sometimes hotly contested), with no scholarly consensus on a definition for the term “gender” (Muehlenhard & Peterson, 2011) or “gender identity” (Bussey, 2011). Dozens of major theories of gender development

have been proposed over the years (e.g. Freud, 1905/1962; Kohlberg, 1966; Kessler & McKenna, 1978; Eagly, 1987; Buss, 1995), and each approach has described new gender-related constructs and produced new measures to study those constructs. Noting that Carole Beere's 1990 review identified over 1,400 gender-related measures, Smiler and Epstein (2010) observed that "the sheer volume of measures suggests...that gender is a complex, multidimensional phenomenon" (p. 133). Indeed, several scholars have explicitly put forward multidimensional models of gender (e.g. Bussey & Bandura, 1999; Egan & Perry, 2001; Fausto-Sterling, 2012; Tate et al., 2014; van Anders, 2016).

In spite of the evidence that both gender and sex are multidimensional (e.g. Diamond et al., 2011; Fausto-Sterling, 2012), the culturally omnipresent idea that gender is determined by genitals leads most researchers to assume that asking for participants' genders will tell them something about those participants' genitals and/or other aspects of their physical bodies. That is, they continue to behave as though gender and sex are dichotomous and perfectly linked. It is this assumption which underlies the ubiquitous generalizations from a person's assertion that they are a man or a woman (i.e. a binary gender question) to scientific claims about their anatomy, hormonal status, fertility, appearance, social role, and many other characteristics.

Given the existence of transgender people, this approach is clearly flawed. By definition, transgender people assert gender identities other than the ones presumed to follow from their genitals and/or socialization, and often change their bodies and social roles to better reflect their gender identities (e.g. Diamond et al., 2011). Crucially, from this perspective gender identity trumps both the physical body and socialization, and thus many transgender people will report their gender identity rather than their genital status wherever possible (e.g. Balarajan et al., 2011; Fenway Institute, 2013; Conron et al., 2014). Regardless of whether the question asks about sex

or gender, it is likely that cisgender and transgender women will select “female” while cisgender and transgender men will select “male.” When participants and researchers interpret a question differently (as is the case here), the validity of any research based on those data may be substantially compromised.

As a result, transgender people are less likely to be subject to inappropriate generalizations from a single question in this way (cf. Zucker et al., 2012), as evidenced by the fact that information such as sex assigned at birth, transition history, and gender presentation are commonly measured alongside gender identity whenever transgender people are studied, and reported (or asked about) whenever transgender people’s genders are discussed.<sup>4</sup> Many researchers (quite reasonably) have concluded that it is not acceptable to generalize from a transgender person’s identity as a man or a woman to make scientific claims about their anatomy, hormonal status, fertility, appearance, social role, or any other characteristics.

Unfortunately, rather than continuing this line of reasoning to conclude that these various constructs should be measured separately for *all* participants (whether transgender or cisgender) when they may be relevant to outcomes of interest, many researchers seem to have concluded that it is more efficient to measure *transgender status* (i.e. to distinguish cisgender and transgender people). This reasoning supposes that if cisgender people can be assumed to have unidimensional genders and sexes (with gender having a perfect relationship to sex), no further investigation of multidimensionality is necessary after a person is identified as cisgender.

Transgender people, in contrast, can then be excluded from the sample altogether (in the service

---

<sup>4</sup>However, it is important to note that transgender people’s genders frequently continue to be described primarily in relation to their sex assigned at birth in the psychological literature, regardless of their gender identity, presentation, or social role (Ansara & Hegarty, 2012), an approach which is both cissexist (Serano, 2007), and contrary to the recommendations of the American Psychological Association’s Task Force on Gender Identity and Gender Variance (APA, 2009) and its publication manual (APA 2010, p. 74).

of simplicity), or subjected to an extensive list of questions to assess the various dimensions of their genders and sexes, all of which must be considered in any analysis of their data.

Approaches like these are flawed both because of the data they fail to collect from putatively cisgender participants, and because of the undue scrutiny they place on transgender participants.

Binary gender questions do not permit researchers to distinguish transgender participants from cisgender participants (categories whose definition is not universally agreed-upon in any case). These questions also presume (and thus find “evidence” for) the existence of two and only two genders, while other types of questions find evidence for multiple genders (e.g. Factor & Rothblum, 2008; Johnson & Wassersug, 2010; Kuper et al., 2012), and and/or a gender spectrum (e.g. Joel et al., 2014) among the general population. In a survey of cisgender people (i.e. who self-reported congruent sex assigned at birth and gender identity), 35% identified to some extent as both a man and a woman, neither a man nor a woman, or as a gender incongruent with their sex assigned at birth, suggesting that gender identity more closely resembles a spectrum between identifying as a man or a woman than it does a binary categorization (Joel et al., 2014).

In addition to providing inaccurate data, binary gender questions are disliked among participants who do not identify as men or women (Staples, Bird, Masters, & George, 2018), and they may reduce self-esteem and increase negative affect and feelings of threat in these groups. Townsend, Markus, and Bergsieker (2009) suggest that mixed-race participants who are required to choose only one racial identity on survey questions (and thus are forced to miscategorize themselves) may experience categorization threat, resulting in lowered motivation and self-esteem. To the extent that participants feel they cannot adequately represent their experiences and identities when answering a particular question, or that a question is inappropriate or invasive, they may feel more negatively toward researchers and believe that researchers feel less

positively toward them and toward gender minorities. Indeed, Meier and Labuski (2013) argue that transgender people may stop participating in research which they view as exclusionary, and may warn the rest of the transgender community to avoid a particular research project or researcher.

Pragmatically, these issues are of greatest concern to public health researchers (who, not coincidentally, are perhaps the most vocal and influential opponents of binary gender questions). As a community, transgender people (especially people of color) face very serious health disparities in the United States (e.g. Grant et al., 2011; Institute of Medicine, 2011) and around the world (e.g. Operario et al., 2008; Reisner et al., 2013). Given these growing public health concerns, the issue of how to identify this population in the first place looms large; research and interventions with transgender people cannot proceed without first identifying who is transgender, and binary gender questions make such identification impossible. Developing a standard set of gender questions which are capable of identifying transgender people is therefore a top priority in improving the health of this community. A report by the Institute of Medicine (2011) noted that “one of the greatest challenges to synthesizing scientific knowledge about the health of sexual and gender minorities has been the lack of standardized measures in federal surveys” (p. 303-304). Similar concerns have been raised by the American Public Health Association (1999), American Psychological Association (APA) Task Force on Gender Identity and Gender Variance (2009), and the U.S. Department of Health and Human Services (DHHS, 2013), and each of these organizations has called for a standard set of questions to identify transgender people. Demographers Durso and Gates (2013) neatly summarize the rationale in the following way:

“Discussions of civil rights, program evaluation, public health, and the delivery of human services must rely on sound facts and analyses that come at least in part from high-quality



survey research. Facts about sexual minorities are often not available because specific questions pertaining to sexual orientation and gender identity are not routinely included on surveys, leaving scholars, policymakers, and the general public to risk falling back on stereotypes and myths about the experiences and social situations of LGBT people” (p. 21-22).

Clearly, both academics and the transgender community have much to gain from the development and implementation of better alternatives to the binary gender question. From a public health perspective, research on the health outcomes of transgender people is desperately needed, and that research cannot proceed without a basic set of questions capable of identifying transgender people. The advancement of theories of gender development and difference is likewise hampered if researchers exclude or overlook transgender experiences, or routinely collect imprecise or incomplete data. Without theory-driven specificity in their measurement of gender, researchers cannot say *which* gendered constructs are driving any observed gender differences, which severely limits their ability to understand the causes, correlates, and outcomes of these effects. For all of these reasons, it is crucial to develop inclusive alternatives to the binary gender question.

What information should these alternative measures collect? A full consideration of all the gender-related constructs that might be significant to any researcher in any field would be virtually impossible to give (Smiler & Epstein, 2010). The more general point to take away is that any measure of gender or sex can be effective only if it actually asks about the information that is most pertinent to the topic under investigation. While the relevant information may vary dramatically over time and between researchers and disciplines, it is essential for every researcher to take the time to carefully consider what information they really need and for what purpose. As I am a social psychologist, I focus on identifying successful measures of *gender*

*identity*, a dimension of gender with special relevance for social psychology, and *transgender status*, an understudied intersecting identity.

### **Theoretical Perspectives on Gender Identity**

Gender identity as a distinct construct has mostly been described in relation to children, intersex people, and transgender people to distinguish the subjective sense of one's own gender membership from the gender membership socially assumed to follow from one's genitals. Authors from various disciplines have produced often conflicting descriptions of gender identity as innate or socially constructed, unchanging or malleable, each based in different (and often nearly unfalsifiable) assumptions about the nature of an identity more generally. Of course, these various definitions usually lead to the same conclusions about the gender identities of the majority population of adult cisgender people, and until very recently, theories of gender identity development have attempted to explain only the development of a cisgender identity (e.g. Kohlberg, 1966; Martin & Halverson, 1981), leaving the development of transgender identities under the purview of clinical psychology (e.g. Zucker, 2005; deVries et al., 2014). While a few authors have proposed unifying theories of gender identity development (e.g. Diamond et al., 2011; Fausto-Sterling, 2012), their theories have not yet generated much change in the study of gender identity. For this reason, even scholars of gender development have frequently presumed that the distinction between genitals and gender identity is mostly theoretical, and that given enough time, particular genital configurations almost invariably lead to a sense of oneself as a man or woman.

Empirically, most psychologists (e.g. Kohlberg, 1966; Martin & Halverson, 1981) have defined gender identity as the ability to label oneself and others as a boy or girl. This ability has nearly always been described as based in either innate or social knowledge of one's "biological

sex,” and presumes the existence of two mutually exclusive gender categories. For instance, Spence and Buckner (1995) argued that “children’s recognition of their biological sex is almost invariably accompanied by the development of what has been called gender identity, a basic existential sense and acceptance of themselves as male or female” (p. 115). Other authors (e.g. Money, 1955) have argued that gender identity is merely a product of assignment to a category. In this view, simply being told that one is a boy or a girl is sufficient to produce a gender identity as a boy or a girl.

Obviously, these theories (e.g. Kohlberg, 1966; Martin & Halverson, 1981) explain only the development of a binary cisgender<sup>5</sup> identity, failing to explain how or why transgender children develop a gender identity that is incongruent with their assigned sex at birth, or how some people develop non-binary<sup>6</sup> gender identities. Many gender theorists and social scientists (e.g. Kessler & McKenna, 1978; Butler, 1990; Bornstein, 1994, Califia, 1997; Diamond, 2000; Fausto-Sterling, 2012) have strongly critiqued this tendency to equate gender identity (in the sense of category membership) with genitals. Citing the experiences of transgender people, these authors argue that a person’s sense of their gender is not determined by their physical body or by the socialization they receive from others. Rather, they argue that gender’s essence lies in a separate inborn knowledge of self, which may mean feeling like a man, a woman, or something else entirely. Crucially, this perspective asserts that gender identity trumps both the physical body and socialization, observing that transgender people assert gender identities other than the ones

---

<sup>5</sup> I define cisgender as “a person whose gender identity is congruent with the one socially expected based on their sex assigned at birth.” In Chapter 3, I problematize this term, asking whether intersex people, non-binary people, agender people, and so forth should be considered “cisgender” or “transgender” and exploring the consequences of each choice.

<sup>6</sup> I define non-binary gender identities as “gender identities which are not male or female” (e.g. Factor & Rothblum, 2008; Johnson & Wassersug, 2010; Kuper et al., 2012).

presumed to follow from their genitals and/or socialization, and often change their bodies and social roles to better reflect their gender identities (e.g. Diamond et al., 2011).

Other authors concerned specifically with the development of transgender identities have argued variously that gender identity in transgender people is an essential, innate, unchanging characteristic (e.g. Ehrensaft, 2012), that gender identity is an unstable characteristic predisposed by biology and influenced by social factors (e.g. Zucker et al., 2012), or even that “there is no gender identity behind the expressions of gender; that identity is performatively constituted by the very ‘expressions’ that are said to be its results” (Butler 1990, p. 25). Unfortunately, these perspectives also fail to explain why a particular person develops a particular identity. Given that virtually all children are born with genitals and are socially assigned to a male or female category, why do some children develop along cisgender trajectories while others develop along transgender ones? No substantive answer to that question has yet been found, but accumulating research does make clear that multiple factors influence gender identity (deVries et al., 2014).

One important recent insight is that individuals play active roles in constructing their own gender identities over the course of their lives (Bussey & Bandura, 1999; Diamond et al., 2011; Fausto-Sterling, 2012). That is, a person’s gender identity may shape the environment around them even as it is shaped by that environment. This reciprocal relationship between personal, social, and environmental factors is at the core of two of the most recent models of gender identity development, both of which attempt to integrate our knowledge about transgender identity with theories of cisgender development to produce a theory which can explain both identity trajectories.

One early model of gender identity in cisgender people was proposed by Bussey and Bandura (1999), who argued that gender has multiple determinants, (including personal, social, behavioral, and environmental factors) which reciprocally influence one another to produce what we describe as gender. They further asserted that people are agents who actively contribute to their own social development (for instance, by seeking out particular social or environmental contexts), rather than merely the passive products of external influences.

A similar argument was later advanced to describe gender identity in transgender people. Although transgender identity development has often been conceptualized as a clear, linear process of bringing one's physical body into alignment with an already-existing gender identity (e.g. Devor, 2004), Diamond et al. (2011) argue lucidly that this trajectory does not describe all transgender experiences. Rather, they draw on dynamical systems theory to suggest that gender identity development may for some be a recursive process that changes over time, with no clear endpoint. They also reject the view that gender identity is an innate, unchanging characteristic of an individual, arguing that it is instead "a hard-fought achievement, a truly novel creation forged out of the individual's entire history of gender experience and his/her creative explorations of new forms of gendered self-expression" (p. 638). This argument highlights the possibility that a given person's gender identity may change and evolve over time in response to social, individual, and contextual factors.

Anne Fausto-Sterling (2012) takes a similar view of gender identity in both transgender and cisgender people, arguing that gender identity should be viewed as a dynamical system, or a collection of systems changing over time. Like Diamond et al. (2011), she argues that gender identity can never be fixed but only stable, and that all people must go through a process of developing a gender identity. Critiquing simplistic views of biology as a mere catalyst or a

unidirectional force shaping gender, she asserts that the body is “not the foundation of all things, but rather is in the middle - sustained within the world, responding to it, but also reshaping it” (p. 404). She further notes the importance of cultural context and intersectionality, recognizing that the exact factors which influence gender identity development will differ across contexts.

In sum, several contemporary theorists suggest that gender identity forms through the reciprocal influence of a variety of factors, and that it develops over time in both cisgender and transgender people. Rather than having an essential, unchanging core, gender identity is best viewed as a complex system which may be stable or change over time.

### **Multidimensional and Integrative Measures of Gender Identity**

The literature on gender identity’s origins strongly suggests that it is a complex phenomenon, but very few researchers have attempted to measure gender identity in any detail. Indeed, as described at the beginning of this paper, the vast majority of researchers persist in asking about gender as a single construct, so it is hardly surprising that gender identity has received less attention. However, a few researchers have developed multidimensional gender measures, and I review them here.

Egan and Perry (2001) designed a multidimensional gender measure explicitly for use with children. Their scale had five dimensions: knowledge of which gender category one belongs to, feelings of typicality for that gender and contentedness with that gender assignment, felt pressure to behave according to gender stereotypes, and bias toward one’s own gender group. Children’s self-esteem and peer acceptance were affected by each of these dimensions, though the dimensions themselves are not highly correlated. This measure is not without flaws, however; it groups children by their sex assigned at birth, and aims to determine whether a given child is “typical” of children with the same sex assigned at birth. It further fails to acknowledge

the possibility of non-cisgender gender outcomes. In their model, there are only conforming and non-conforming members of binary, pre-determined assigned sex at birth categories.

Tate et al. (2014) both build on and critique the work done by Egan and Perry (2001) with their own “gender bundle” multidimensional model of gender identity. While their five dimensions (birth assigned gender category, current gender identity, gender roles and expectations, gender social presentation, and gender evaluations) differ slightly from those of previous authors, there are enough similarities that Tate et al. (2014) explicitly compare their model to that of Egan and Perry (2001) point-by-point. However, Tate et al. take a decidedly different perspective on the appropriate point of comparison for a gender identity question to work from. Rather than using sex assigned at birth as a reference point, they group participants by gender identity, pointing out that there is “no evidence of asymmetry across transgender and cisgender actors” with respect to gender development. That is, there is no evidence to support the notion that transgender people’s gender identities are any less predictive of outcome variables than are the gender identities of cisgender people (Tate et al., 2014). When the topic of study is social psychological variables, it therefore seems illogical to combine participants on the basis of their assigned sex rather than their shared sense of themselves as members of a particular gender category. Instead, transgender people might be more accurately and parsimoniously combined with the conforming members of the other assigned sex on the basis of their shared gender identity than labeled as highly non-conforming members of a category based on assigned sex at birth. The deliberate erasure of transgender identities in Egan and Perry’s measure is thus both theoretically inadequate and ethically dubious, as it perpetuates cisgenderism.<sup>7</sup>

---

<sup>7</sup> Cisgenderism in psychology causes transgender people’s gender identities to be viewed as invisible, delusional, or pathological, resulting in numerous theories of gender development and expression which implicitly assume a cisgender outcome and discount transgender people’s descriptions of themselves (e.g. Bauer et al., 2009; Ansara & Hegarty, 2012).

Tate et al. (2014) suggest several reasons why psychologists might have skewed perceptions of transgender people's gender self-categorization. For instance, the frequent use of terms such as "male to female" or "MTF" put excessive emphasis on a presumed former gender category. However, there is no evidence that transgender children's gender identities were ever congruent with their sex assigned at birth, and substantial anecdotal and empirical evidence in support of the idea that transgender children's gender identities are often present and stable early on in development. Thus, while a child's development will certainly be influenced by the assumptions others make about their genders, those external assumptions are not the same as an identity. A second reason for inaccurate assumptions about transgender people's self-categorization may be the tendency for transgender narratives to be ignored, dismissed, or contested. Despite the wealth of personal stories written by transgender people (e.g. Morris, 1974; Bornstein, 1994), psychologists have tended to dispute or ignore these subjective experiences by dismissing them as the products of a psychological disorder. With all individuals who report experiences of gender which differ from the expected model thus pathologized, all data points which might otherwise change the model are disregarded (e.g. Bauer et al., 2009).

One recent paper has provided some additional empirical evidence to support Tate et al.'s argument that cisgender and transgender people's gender identities are the same construct. Olson et al. (2015) carried out another comparative study of cisgender and transgender people in an attempt to refute claims that transgender children are merely pretending to be their expressed gender or are simply being oppositional. They gave both explicit (gender identity, gender peer preferences, and object preference) and implicit (a gender-identity and gender-preference IAT) measures of gender identity to 32 prepubescent transgender children (aged 5-12), all of whom were highly conforming to their gender identity and were perceived as their gender identity in



daily life. The same measures were given to their siblings, and control participants matched for gender identity and major demographic characteristics. Transgender children consistently identified with their gender identities on both explicit and implicit measures, and in this respect were not significantly different from cisgender children with the same gender identities. These findings suggest that the gender identities of transgender people are as deeply-held and empirically predictive as those of cisgender people.

I define *gender identity* as the relationship a person perceives between the self and the gender groups commonly recognized within their culture. While its precise origins are not well understood, the best available evidence suggests that gender identity is a complex dynamical system reciprocally influenced by biological and psychological characteristics of the individual and by social and environmental factors. Thus, gender identity may be stable or may change over time (e.g. Diamond et al., 2011; Fausto-Sterling, 2012). A person's gender identity is not necessarily the same as external observers' perceptions of their gender group membership (e.g. Tate et al., 2014). Rather, I argue that like other social identities, it is quintessentially based in *self-categorization*, a subjective sense of membership or lack of membership in a given gender group (e.g. Turner, 1982). This is consistent with previous conceptualizations of gender identity as a sense of one's own relationship to existing gender groups (e.g. Egan & Perry, 2001; Tate et al., 2014). In hopes of countering the cisgenderist trend in previous research and advancing the integrative study of cisgender and transgender people's gender identities, I also explicitly view gender identity as the same construct for transgender and cisgender people, as well as for people with binary gender identities and non-binary gender identities (e.g. Diamond et al., 2011; Fausto-Sterling, 2012; Tate et al., 2014; Olson et al., 2015).

### **Defining and Measuring Transgender Status**

While developing theoretically precise measures of *gender identity* that accommodate both cisgender and transgender participants is an important goal for social scientists, it is equally important to develop theories and measures of *transgender status* that can capture the distinct life experiences of cisgender and transgender people. However, it is far from straightforward to operationally define “transgender.” As with gender identity, transgender experience has multiple dimensions which are emphasized to varying degrees by different researchers and by transgender people themselves. The following section describes four common ways that researchers have defined the category “transgender,” and the benefits and limitations of each approach.

### **Measuring gender identity and sex assigned at birth.**

As discussed earlier, incongruence between sex assigned at birth and gender identity is widely accepted by psychologists as the defining characteristic of transgender people (e.g. American Psychological Association, 2009; American Psychiatric Association, 2013; Institute of Medicine, 2011). Cisgender people are those whose sex assigned at birth is the same as their current gender identity, and transgender people are those whose sex assigned at birth differs from their current gender identity.

Many researchers seem to feel that asking participants for their sex assigned at birth will provide information about their physical bodies. However, as several qualitative studies have revealed, participants interpret the question to mean “whether the doctor decided you were a boy or a girl when you were born,” (e.g. Lombardi et al., 2013), which is almost entirely determined by the appearance of an infant’s external genitalia at birth (e.g. Hubbard, 1996). Contrary to popular assumption, then, asking about “sex assigned at birth” does not consistently elicit accurate information about participants’ physical bodies at birth, their current physical bodies, or the sex listed on their birth certificate.

Sex assigned at birth is not equivalent to a person's physical body at birth; even obviously intersex babies are assigned male or female at birth in most countries (e.g. Consortium on the Management of Disorders of Sex Development, 2006). It is likewise not equivalent to a person's current physical body, as many people change aspects of their anatomy and hormonal makeup in their lifetimes as a result of ordinary aging processes or medical intervention. Thus, asking only about sex assigned at birth clearly does not provide sufficiently detailed information about a person's physical body for use in medical or public health research. For this reason, WPATH (Deutsch et al., 2013) recommended the use of an "anatomical inventory" (i.e. a multiple-choice list of organs that a person might have) to provide doctors with sufficiently specific information for the provision of basic medical care to transgender patients.

Sex assigned at birth also does not reliably provide information about the sex listed on a person's birth certificate, as that information can be legally changed in many U.S. states (e.g. Grant et al., 2011). There are several components of legal sex, including the sex listed on a person's birth certificate, drivers' license, passport, social security card, and various insurance policies, and for many transgender people these components do not match (e.g. Grant et al., 2011). Asking about sex assigned at birth cannot provide accurate information about each of these components of legal sex. Given that consistency of documentation can have substantial effects on a person's access to housing, transportation, employment, and health care, measuring those components of legal sex separately may be crucial to fully understanding their situation.

### **Measuring history of transition.**

Another possible definition of "transgender" is "having a history of transition." Early studies of transgender people were conducted almost exclusively through gender clinics providing medical transition services (e.g. Lawrence, 2010; Meier & Labuski, 2013), and this

view of the category “transgender” has in some cases persisted into the present. For example, Lombardi (2009) asked participants to disclose whether they were currently using hormones or had undergone top or bottom surgery, among other questions. Xavier et al. (2007) also defined transgender participants (among other criteria) as those who had or wished to physically modify their bodies to match their identities.

This definition of “transgender” is perhaps the most physically centered. While the wording used in each question varies, it often implicitly or explicitly focuses on physical aspects of transition (e.g. hormones, surgeries) rather than self-identity. As such, this question may feel invasive to participants who are uncomfortable disclosing physical information about themselves. It may also (depending on the wording) omit participants who wish to undergo transition but have not yet, those who have undergone some but not other aspects of transition (as there are a wide variety of possible interventions to choose from; e.g. Coleman et al., 2012), and those who have undergone transition but conceptualize it in different terms (e.g. gender affirmation surgery as opposed to gender reassignment surgery have different implications about the source of a person’s gender). Its benefits include the ability to include people who may not identify themselves as transgender (see the following section). It may also be appropriate for medical contexts because of its focus on biological rather than social factors.

### **Measuring transgender identity.**

Some researchers studying transgender people have been primarily interested in whether a person self-defines as a member of the social category “transgender.” Many local and national transgender needs assessments have used this sort of definition (e.g. Grant et al., 2011, Conron et al., 2012, Conron et al., 2014). For example, the National Transgender Discrimination Survey (Grant et al., 2011) used self-definition as a transgender person as their sole exclusion criterion.

Likewise, Conron et al. (2012) asked whether their participants think of themselves as transgender. This type of question also bears a striking similarity to the single-item measure of social identification proposed by Postmes, Haslam, and Jans (2013) for general use in assessing self-categorization.

While an identity-based approach may appear to be the most inclusive way to define “transgender,” it seems likely that a question about transgender identity alone will overlook people who have a history of transition but do not identify as transgender. For instance, noting that qualitative research about transgender identities has frequently distinguished between openly transgender people and people who do not disclose their history of transition (i.e. “stealth” transgender people), Tate et al. (2013) suggest that openly transgender people may select options with the word “transgender” in them (e.g. “transgender woman”), while “stealth” transgender people may prefer to select an option without the “transgender” modifier (e.g. “woman”). This general premise is consistent with the idea that some people with a history of transition may not identify as transgender. In addition, people who self-describe as “transgender” may have qualitatively different experiences of life than those who do not self-describe in this way (e.g. different levels of concealment of a history of transition).

### **Measuring gender presentation.**

Gender presentation, or the degree of one’s conformity to stereotypical behaviors and appearances for socially recognized gender categories, is another way to operationally define transgender identity. This view emphasizes a person’s social conformity to a gender group over self-identity or the person’s physical body, and tends to include a broader array of gender non-conforming, non-transitioning, cisgender-identified people. Researchers taking this approach have largely used a measure of socially assigned gender nonconformity developed by Wylie et

al. (2010), which includes questions about the femininity and masculinity of a person's appearance and mannerisms. Where a person's developmental history is of interest, or where converging evidence from close family members is desired, some researchers have also used the Recalled Childhood Gender Identity/Gender Role Questionnaire (Zucker et al., 2006), which asks a series of questions about the various gendered activities, behaviors, and relationships that the person engaged in as a child. Both scales ultimately compare this information to a person's sex assigned at birth.

Several studies have found that gender nonconformity seems to predict experiences of discrimination regardless of sexual orientation or transgender status, although LGBT people are at higher risk (e.g. Horn, 2007; Blashill & Powlishta, 2009; Stotzer, 2009). Wylie et al. (2010) further observe that this relationship may be U-shaped, as transgender people who are highly nonconforming relative to their birth sex are more likely to pass as the gender with which they identify (and thus to escape discrimination). Thus, the inclusion of this measure with other measures of gender-related constructs may help researchers to distinguish subgroups of transgender people who are particularly at risk for discrimination.

However, this type of measure should also only be included if researchers have a clearly defined theoretical reason to request it. In cognitive testing, Lombardi et al. (2013) found that a majority of both cisgender and transgender participants indicated conformity to their gender identity, limiting its usefulness in distinguishing subgroups. They also noted that this type of question may cause distress for gender minority participants, who often recall unpleasant experiences associated with not "passing" as cisgender members of their gender identity. It also seems likely that the inclusion of this measure will create feelings of pressure for transgender

participants to “prove” their identities by conforming highly to stereotypes for their gender identity.

### **Conclusion**

Binary gender questions are widely used but have been criticized for their inability to identify gender minorities (transgender and gender non-conforming people) and their inconsistency with modern multidimensional theories of gender. While substantial research suggests that gender is multidimensional (e.g. Diamond et al., 2011; Fausto-Sterling, 2012), very few validated gender measures incorporate this insight. Many current theories and measures of gender also typically overlook transgender and non-binary people, limiting the ability of researchers to learn about these populations or to develop robust models of gender development and difference.

I propose that researchers should ask about gender identity and transgender status separately, and should use precise definitions of each construct they measure. Distinguishing between sex assigned at birth, having a history of transition, and identifying as transgender may enable researchers to learn more about subgroups within the transgender community and develop a more sophisticated understanding of transgender people’s lives. Membership in these each of these categories may predict outcome variables differently. In any event, awareness of the distinction between these concepts will help researchers more accurately assess how the data from particular types of questions can be appropriately generalized.

## **Chapter 2: Alternatives to the Binary Gender Question**

The inadequacies of the binary gender question have been recognized for at least a decade, and numerous researchers and organizations have produced their own novel gender questions. The vast majority of existing research on alternatives to binary gender questions has been conducted by population-level, public health, and medical researchers with the explicit goal of identifying transgender people and/or distinguishing transgender people from cisgender people. These researchers have therefore generally drawn a strong distinction between physical and psychological aspects of gender, as a “mismatch” between these two is widely accepted as the defining characteristic of transgender people (e.g. APA, 2009; Institute of Medicine, 2011). The result of this focus has been a two-question method, which asks about gender identity and transgender status separately. By far the most commonly used of these methods is one which asks about current gender identity and sex assigned at birth in order to classify participants as cisgender (if sex assigned at birth is the same as current gender identity) or transgender (if sex assigned at birth differs from current gender identity).

This pair of questions (gender identity + sex assigned at birth) has seen relatively wide acceptance, having been recommended for use in clinical settings and electronic health records (Sausa et al., 2009; Deutsch et al. 2013; Fenway Institute, 2013; Cahill & Makadon, 2014), population-level public health and demographic surveys (GenIUSS Group, 2014; Reisner et al., 2014), and general research (Tate et al., 2013). This measure is a substantial improvement over the binary gender question in that it recognizes a distinction between physical body and identity, and allows researchers to distinguish transgender and cisgender participants. In addition, certain formulations of the question enable researchers to identify participants who identify as something other than men or women. These improvements have the potential to dramatically



improve the quality and availability of research about transgender communities and about gender variance more generally, and for the reasons previously explained, these changes are sorely needed.

However, the development of this measure (and most other inclusive gender questions) has proceeded mostly without the benefit of a clear theoretical rationale. Most questions have instead reflected the (largely unsupported) preconceptions of their creators (e.g. Sausa et al., 2009), sometimes modified to accommodate the opinions of focus groups of transgender and cisgender respondents (e.g. Balarajan et al., 2011; Lombardi et al., 2013; Conron et al., 2014). Perhaps as a result, the precise wording of the questions and response options vary (sometimes dramatically) between organizations and researchers, often with little justification. In combination, these factors mean that careful examination of the wording used by each researcher is necessary to draw conclusions about the efficacy of each method. This is the aim of the following sections. Please refer to Table 1 for the full text of each question.

Table 1

*Full Text of Previously-Published Gender Identity and Sex Assigned at Birth Measures (“Two-Question Methods”)*

Sex Assigned at Birth Question	Gender Identity Question
<b>Kenagy (2005)</b>	
<i>What was your physical sex at birth?</i>	<i>What is your gender identity?</i>
Male	Male to female transsexual
Female	Female to male transsexual
Intersex	Transman
	Transwoman
	Passing butch
	Transvestite
	Intersexed
	Drag king
	Drag queen
	Cross dresser
	Male
	Female
	Transgendered
	Transgenderist
	Other [free-text]
<b>Boston Gay, Lesbian, Bisexual, and Transgender Health Access Project (GLBT HAP; Sperber, Landers, &amp; Lawrence, 2005)</b>	
<i>What is your sex or gender?</i>	<i>What is your transsexual/transgender identity?</i>
Transgender	Male to Female
Transsexual	Female to Male
Male	Trans
Female	Male
Other [free-text]	Female
	Other [free-text]
<b>Center of Excellence for Transgender Health (Sausa et al., 2009)</b>	
<i>What sex were you assigned at birth? (Check one)</i>	<i>What is your sex or gender? (Check all that apply)</i>
Male	Male
Female	Female
Unknown or Question Not Asked	Transgender Male/Transman
Decline to State	Transgender Female/Transwoman
	Genderqueer
	Additional Sex or Gender [free-text]
	Unknown or Question Not Asked
	Decline to State

Sex Assigned at Birth Question	Gender Identity Question
<b>National Transgender Discrimination Survey (Grant et al., 2011)</b>	
<i>What sex were you assigned at birth, on your original birth certificate?</i>	<i>What is your primary gender identity today?</i>
Male	Male/Man
Female	Female/Woman
	Part time as one gender, part time as another
	A gender not listed here, please specify [free-text]
<b>Equality and Human Rights Commission (Balarajan et al., 2012)</b>	
<i>At birth, were you described as.... (Please tick one option)</i>	<i>Which of the following describes how you think of yourself? (Please tick one option)</i>
Male	Male
Female	Female
Intersex	In another way [free-text]
I prefer not to say	
<b>National Center for Health Statistics (Redford &amp; Van Wagenen, 2012)</b>	
<i>What was your sex at birth?</i>	<i>Do you consider yourself to be male or female?</i>
Male	Male
Female	Female
Intersex/Ambiguous	
<b>Lombardi et al. (2013)</b>	
<i>What sex were you assigned at birth on your original birth certificate? (Check one)</i>	<i>What is your sex or gender? (Check ALL that apply)</i>
Male	Male
Female	Female
Unknown or Question Not Asked	Other: Please specify [free-text]
Decline to State	
<b>The Fenway Institute (2013)</b>	
<i>What sex were you assigned at birth on your original birth certificate? (Check one)</i>	<i>What is your current gender identity? (Check all that apply)</i>
Male	Male
Female	Female
Decline to Answer, please explain why	Female-to-Male (FTM)/Transgender
	Male/Trans Man
	Male-to-Female (MTF)/Transgender
	Female/Trans Woman
	Genderqueer, neither exclusively male nor female
	Additional Gender Category/(or Other), please specify

Decline to Answer, please explain why

---

---

Sex Assigned at Birth Question	Gender Identity Question
<b>Tate et al. (2013)</b>	
<i>What gender were you assigned at birth?</i>	<i>What is your current gender identity?</i>
Female	Female
Male	Male
Intersex	Transgender
	Genderqueer
	Intersex
<b>The Massachusetts Gender Measures Project (Conron et al., 2014)</b>	
<i>What is your sex?</i>	<i>What is your gender?</i>
Male	Male
Female	Female
	Transgender
<b>Reisner et al. (2014)</b>	
<i>What sex were you assigned at birth, on your original birth certificate? (check one)</i>	<i>How do you describe yourself?</i>
Female	Female
Male	Male
	Transgender
	Do not identify as female, male, or transgender

---

## **Previously Proposed Inclusive Gender Questions**

Most of the first inclusive gender questions were developed by groups of transgender activists. For instance, one of the earliest versions was developed in 1997 by a Philadelphia transgender health advocacy group (Kenagy, 2005), and was then used in several other regional studies (Xavier, 2000; Xavier et al., 2007). In 2009, a modified version of this method was recommended by UCSF's Center of Excellence for Transgender Health (Sausa et al., 2009), and due to its status as a major transgender health advocacy organization, this general design was subsequently adopted by other researchers and organizations with an interest in transgender health (e.g. Institute of Medicine, 2011; Deutsch et al. 2013; Fenway Institute, 2013; Cahill & Makadon, 2014; GenIUSS Group, 2014; Reisner et al., 2014).

### **Preliminary Validity Evidence: Lombardi et al. (2013), Tate, Ledbetter, & Youseff (2013), and Fenway Institute (2013).**

Among the first U.S.-based researchers to study the validity of a two-question method were Lombardi et al. (2013), who carried out cognitive testing on a slightly simplified version of the two-question measure recommended by Sausa et al. (2009) with a Midwest convenience sample of 25 cisgender and 25 transgender/gender-nonconforming individuals. While Lombardi et al. (2013) argue that their question pair was “easy to use and understood by both gender minority and majority populations” (p. 21), they also report that many gender minority participants found the question “What is your sex or gender?” confusing. The vast majority of participants regarded sex and gender as two separate categories, with “sex” being biological and “gender” being social/psychological. This led many transgender participants to select two options for the first question; one to reflect biological characteristics and the other to reflect identity. For these reasons, Lombardi et al. recommend that in future research the first question

should ask only “What is your gender?” rather than asking about both sex and gender. This modified question wording is also reflected in recent recommendations published by the Fenway Institute (2013), the World Professional Association for Transgender Health (Deutsch et al., 2013), and by various independent medical researchers (e.g. Cahill & Makadon, 2014); however, none of these authors cite Lombardi et al. when making this change, so their rationale for doing so remains unclear.

Although they seem to have developed their questions independently, Tate et al. (2013) carried out another test of a two-question method which closely resembles the modified version eventually recommended by Lombardi et al. (2013). As scholars in San Francisco, Tate and colleagues noticed that many local researchers attempted to identify transgender participants with a single gender identity question using “male,” “female,” “transgender,” and “other” options. However, Tate et al. reasoned that a two-question method might be preferable to the single-question method because it enables researchers to identify genderqueer and intersex people as well as transgender people, and because it reliably distinguishes cisgender and transgender people in a way that a single question cannot (i.e. because most transgender people identify as men or women). To determine which question was more effective, Tate et al. compared the missing and valid data rates across the two measures in seven samples of undergraduate students and community members from the San Francisco area. They found that the number of valid “transgender” and “other” responses was smaller than the missing data rate for the single question measure. However, when using this question pair in a different sample from the same undergraduate population, they found twice as many transgender responses and had virtually no missing data.

Despite the partial inconsistency of these findings, Tate et al. (2013) recommended their question pair to distinguish cisgender and transgender people for the purposes of health care and basic research. Their work has since been cited as conclusive evidence of the two-question method's effectiveness by several high-profile organizations (e.g. Deutsch et al. 2013, p. 701; GenIUSS Group 2014, p. 9). Indeed, on the basis of Tate et al.'s research alone, the Fenway Institute (2013) stated that "research has indicated that the two-step gender identity and birth sex question design performs extremely well" (p. 10). Yet at best, Tate et al.'s results show only that the two-question method results in a lower rate of missing data and higher rate of transgender spectrum responses than the single-question method.

The only evidence Tate et al. provide for the validity of their classifications comes from Study 2b (Sample 5; N = 130). Noting that three students in this undergraduate human sexuality course had publicly identified themselves as transgender during a group activity earlier in the semester, they argue that "the fact that exactly three individuals from this sample were determined to be on the transgender spectrum of gender identity using the two-question method (at a later date in this anonymous survey) demonstrates the predictive validity of the method" (Tate et al. 2013, p. 772). While this is better than nothing, it is far from a rigorous validation. The matching number does not mean they were necessarily the same three people. In addition, the class may have contained other transgender students who were both unwilling to disclose their transgender status publicly *and* unwilling to disclose their sex assigned at birth to a researcher.

However, some of the Fenway Institute's own research on this question pair suggests that a substantial number of transgender people have concerns about answering questions about their sex assigned at birth when posed by a medical provider. The Fenway Institute researchers

assessed reactions to this question pair by administering a short survey to a gender-diverse sample of 251 patients at four community health centers in the Boston, MA area. Consistent with previous research, they found that 97% of participants answered both of the questions and that a majority of participants understood what the questions were asking. However, a substantial portion (17%) of transgender respondents indicated that they wanted to change the gender identity questions. Many others (14% of trans men and 29% of trans women) stated that they were unsure about answering or would not answer the sex assigned at birth question if it were asked at registration, citing the sensitivity of the information. When asked if this information was important for health providers to know, 20% of trans men and 26% of trans women disagreed. It seems from these results that a substantial minority of transgender people consider sex assigned at birth too sensitive to disclose to their medical providers, and do not see clear reasons for their providers to need that information. The Fenway Institute researchers downplay these potential limitations of the method they propose, instead arguing for the method's widespread adoption by other health care organizations.

Despite numerous claims to the contrary, then, research by Tate et al. (2013), Lombardi et al. (2013), and the Fenway Institute (2013) sheds only minimal light on the efficacy of a two-question pair asking about gender identity and sex assigned at birth. Each of these researchers bases their support for this method primarily on the fact that most participants will answer both questions instead of skipping them. Although both Lombardi et al. (2013) and the Fenway Institute (2013) report that some participants were uncomfortable with the sex assigned at birth questions, neither considers the possibility that some groups of participants are disproportionately likely to skip certain questions or to provide inaccurate information.



**Conflicting Evidence: Conron, Lombardi, and Reisner (2014), and Reisner et al.**

**(2014)**

More conclusive evidence of this method's limitations comes from the Massachusetts Gender Measures Project, which tested several gender-related measures through formative focus groups, cognitive testing, and pilot testing with attendees at a transgender youth conference (Conron et al., 2014). Participants completed a modified measure of gender identity and sex, and a measure of transgender status as well as measures of parental gender pressure, bullying, current gender expression and recalled childhood gender nonconformity. By aggregating this information with qualitative notes taken from focus groups, cognitive testing, or personal communications with the participants, the authors generated a "gold standard classification of transgender status" (Conron et al., 2014, p. 971) against which they assessed the classification accuracy of each individual measure.

They report that

...21.9% of transgender youths (7 of 32) did not provide a valid response to the sex item (four skipped it, another selected both male and female, and two chose responses that were inconsistent with their assigned sex at birth). Thus two-item classification approaches that rely on responses to a common sex measure may not accurately discriminate between trans- and cisgender youths (Conron et al., 2014, pp. 971-972).

In total, Conron et al. (2014) found that nearly 30% of their transgender participants were misclassified when only the questions about gender identity and sex were used, prompting them to call for the development of a valid measure of sex. These findings suggest that at least some formulations of this method may be less valid than researchers have hoped, perhaps leading to the collection of excessive missing and/or inaccurate data.

In stark contrast to these findings, Reisner et al. (2014) report unambiguously positive conclusions about the method on the basis of cognitive testing with a subsample of respondents from the Growing Up Today Study (GUTS). Every two years since 1996, this national longitudinal study has assessed the health of the roughly 16,000 children of female registered nurses participating in the Nurses' Health Study II. Because of the longitudinal nature of the study, the researchers had access to the participants' assigned sex at birth as reported by their mothers at the beginning of the study in 1996, which they compared to participants' self-reported gender identity on a 2010 survey. Compared to cisgender participants, participants who were identified as gender minorities (0.33%) had higher levels of recalled childhood gender nonconformity and socially assigned gender nonconformity, providing evidence for the two-question measure's construct validity. In cognitive testing with a subsample of gender minority participants, all participants self-reported an assigned sex at birth concordant with their maternal-reported assigned sex. While these results may seem promising, the participants in Reisner et al. (2014) are highly unusual, as they and their immediate families have been the subjects of repeated study for most of their lives. It seems plausible that these participants are therefore more comfortable discussing intimate matters with researchers (or at least more resigned to the experience) than the average transgender person might be. Additionally, these participants may have been aware (or could have guessed) that their sex assigned at birth had already been documented by their mothers and was therefore already available to the researchers, diminishing the potential benefits of choosing not to disclose a transgender identity.

Among U.S. researchers, only Conron et al. (2014) and Reisner et al. (2014) have directly tested ability of questions about gender identity and sex assigned at birth to identify transgender participants, and they reach contradictory conclusions. While Reisner et al. (2014) found that this

method correctly classified all transgender participants, Conron et al. (2014) find that it miscategorized nearly 30% of transgender participants. Their studies differ in both question wording (Conron et al. asked participants “What is your *sex*?” while Reisner et al. asked “What is your *sex assigned at birth*?”) and sample (Conron et al. recruited transgender youth conference attendees while Reisner et al. recruited the children of nurses who had already been participating in a longitudinal study for most of their lives). It is possible that Conron et al. obtained such poor categorization rates because the wording “What is your *sex*?” does not acknowledge the coercive nature of sex assignment, a common theme raised by transgender participants in several cognitive interviews on this subject (e.g. Balarajan et al., 2011; Lombardi et al., 2013). Alternatively, Reisner et al. may have obtained perfect categorization only because their participants were well aware that their personal histories were already known to researchers.

#### **Resolving Discrepancies: Balarajan, Gray, and Mitchell (2011).**

Fortunately, the discrepancy between the conclusions reached by Conron et al. and Reisner et al. can be at least partially resolved by considering a multi-part report published two years earlier by the UK-based Equality and Human Rights Commission (Balarajan, Gray, and Mitchell, 2011; Glen & Hurrell, 2012). After reviewing the available literature, Balarajan et al. (2011) developed a list of requirements for a truly inclusive gender question. Finding that no existing questions met their previously established criteria, they drafted a new set of questions which they predicted would be more successful. After revising the questions on the basis of feedback from focus groups of cisgender and transgender participants, the questions went through two rounds of cognitive interviewing (again with both cisgender and transgender participants) and were revised again to produce a final set of five recommended questions. Glen and Hurrell (2012) then gave these exact questions to an online sample of over 10,000

participants who had previously completed the “standard” binary gender question as part of their initial recruitment. This design enabled the researchers to compare participants’ responses to the two types of questions while minimizing the likelihood of order effects.

Comparing responses to gender identity and sex assigned at birth questions alone, Glen and Hurrell identified 0.6% of their sample as gender minorities. However, by analyzing a third question about transition history in addition to these questions, the proportion of gender minority participants more than doubled, to 1.4%. Notably, 1% of participants indicated having a history of transition, and 83% of those participants had identified themselves as only male or only female in the previous method (meaning they would have been categorized as cisgender if questions about gender identity and sex assigned at birth were used alone). On the basis of these findings, Glen and Hurrell concluded that

“there was relatively little overlap between different minority groups, which suggests that, to maximise the opportunities for respondents to choose minority identities, questions 1-3 [sex assigned at birth, gender identity, and transition history] all need to be asked. Relying on only some of those questions would miss some groups of respondents” (p. 9).

Glen and Hurrell’s participants were recruited from the general population of the UK and their question referenced “sex assigned at birth,” reducing concerns that their results were merely an artifact of an atypical sample or an objectionably-worded question. Nonetheless, these researchers still found little overlap between the groups of transgender participants identified by this method and the transgender participants identified by the question about a history of transition. On this basis they concluded that asking about gender identity and sex assigned at birth alone would miss a substantial number of transgender respondents, and that it was therefore crucial to use all three questions.

This work is among the most systematic and theoretically-grounded research on inclusive gender demographic questions to date, but despite its rigor and scope, Balarajan et al. (2011) and Glen and Hurrell (2012) do not seem to have been cited by any gender researchers or organizations working in the United States. As a result, instead of using their theories, designs, and findings to scaffold new research, most U.S.–based researchers have spent the last few years working independently to solve the same problem from scratch, usually with far less institutional support and limited funding. While further research is clearly necessary to fully resolve this issue, the best available evidence at this point suggests that questions about gender identity and sex assigned at birth alone are inadequate for distinguishing transgender participants, contrary to the conclusions of several large organizations.

### **Identifying Questions for Further Study**

Having reviewed alternatives to the binary gender question which have previously been proposed, I then selected several of the most promising inclusive measures for further comparison testing (see Chapters 3, 4, and 5 for results). Three of these questions ask about gender identity alone. As gender identity is an integral part of each measure of this type, a primary question is to determine which response options best optimize for simplicity, validity, and acceptability. The binary gender question is included as a baseline, and is compared to the binary question with an “Other” option and to a multiple-choice gender identity question that includes several more options. I also added two types of free text question (small and large box), as participants in my pilot study expressed strong interest in such a measure. While I do not expect to recommend binary or free text questions, they are included to serve as floor and ceiling measures of simplicity and acceptability. That is, previous research suggests that binary questions maximize simplicity of analysis and minimize acceptability to gender minorities, while

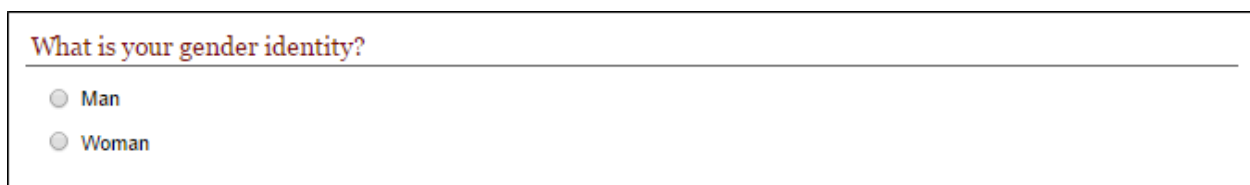
free text questions minimize simplicity of analysis and maximize acceptability to gender minorities. Finally, I included several two-question methods. These combine the same multiple-choice gender identity question with one of three different measures of transgender status (sex assigned at birth, transgender identity, or transition history).

I used the “best-case scenario” version of each question type based on previous research. In all cases, the first question was “What is your gender identity?” both to minimize differences between questions and because the term “gender identity” may cue acceptance of transgender people in a way that “gender” does not. I also consistently used the terms “man” and “woman” as options when asking about gender identity, and the terms “male” and “female” only when asking about sex assigned at birth, as requested by participants in a pilot version of this study. Finally, all questions except the Binary plus Other included “Do not know” and “Choose not to answer” options, following previous research which suggests that this reduces the rate of missing data and makes it easier to tell why people are not answering the question (Redford & Van Wagenen, 2012). In all cases, participants could choose only one option for each question (i.e. questions are single select).

The following sections provide the exact text of the questions used in Study 1 and Study 2 (see Chapters 3, 4, and 5 for methods and results). For each question, I include subheadings indicating the simplicity of analysis, validity, and acceptability of each measure to participants.

### **Measures of Gender Identity.**

#### ***Binary.***

A screenshot of a survey question. The question text is "What is your gender identity?" in a reddish-brown font. Below the question is a horizontal line. Underneath the line are two radio button options: "Man" and "Woman", each preceded by a small grey circle.

*Figure 1. Binary gender identity question used in Studies 1 and 2.*

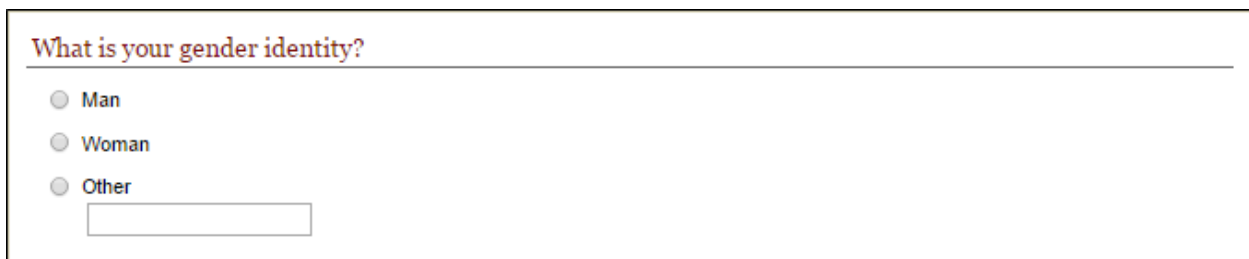
The binary question is included in this study as a sort of control, the status quo against which the other inclusive gender questions can be compared.

**Simplicity of analysis:** As the status quo, this question is by far the most commonly used gender question. It is familiar to participants, very simple to answer, and very simple to analyze.

**Validity:** The binary question lacks validity for most research and practical applications.<sup>8</sup>

**Acceptability to participants:** While this question may not pose much problem for transgender people who identify as men or women, it excludes individuals who do not identify as male or female, forcing people with non-binary gender identities to select one of two inaccurate options. This may induce identity threat for non-binary people (see Chapter 4).

***Binary plus Other.***



What is your gender identity?

Man

Woman

Other

*Figure 2. Binary plus Other gender identity question used in Studies 1 and 2.*

This question makes only a slight modification to the binary gender question, but in theory it includes participants with any gender identity. It is included in this study as a sort of control, as it is perhaps the simplest possible inclusive gender question (i.e. it enables both binary and non-binary people to indicate their genders). It is also becoming more familiar and widely-used in a variety of academic research and by large organizations (e.g. Google asks this questions of users when they set up a Gmail account).

---

<sup>8</sup> For a detailed explanation of the problems with the binary method, refer to Chapter 1.

**Simplicity:** This question is more complex to analyze than the binary question because it includes a free-text box whose contents must be hand-coded. However, given that most people identify as men or women, there are likely to be relatively few of these responses in most samples. This question is also easy for participants to understand, as participants who do identify as men or women can simply ignore the “Other” option.

**Validity:** In theory, this question could be a valid measure of gender identity, because participants can write in any response that is not already provided. However, participants may be uncertain about what researchers mean by “Other,” and so may provide less useful responses. For example, binary transgender people may not select “Man” or “Woman” despite identifying as men or women. Other people may view a free-text box as an opportunity to share sexual orientation, satisfaction with gender identity, or even political views (see Chapter 3) rather than gender identity itself. In practice, validity will be determined more by the coding scheme developed by each researcher to organize the free-text responses than by the question itself.

**Acceptability:** Previous research suggests that binary transgender people may feel identity threat when seeing this question if they imagine that the researcher “expects” them to choose the “other” option because they are transgender (e.g. Pope and Warner, 2016). While this question does allow participants to indicate non-binary genders, it does not explicitly include non-binary people, and so does not provide any cues to indicate how much knowledge the researchers have about non-binary identities.

### ***Multiple-Choice.***

This multiple-choice gender identity question is adapted from one used by the Fenway Institute (2013), which has several features that improve upon the binary + other measure.



What is your gender identity?

Man

Woman

Genderqueer

Another identity not listed (please specify)

Do not know

Choose not to answer

Figure 3. Multiple-choice gender identity question used in Study 1.

Including the term “genderqueer” as its own option was intended to serve as a cue to participants that the researchers who created this question recognize and are interested in genderqueer people. As genderqueer participants make up a substantial proportion of people who do not identify as men or women, including this option may also reduce use of the free-text box and thus the number of responses that need to be coded. The “Decline to answer” option is included to reduce missing data, as many participants who do not wish to reveal this information will otherwise simply skip the question, leaving researchers with no way to know why the information is missing. It is also consistently requested in cognitive interviews with gender minority participants (Balarajan et al., 2011).

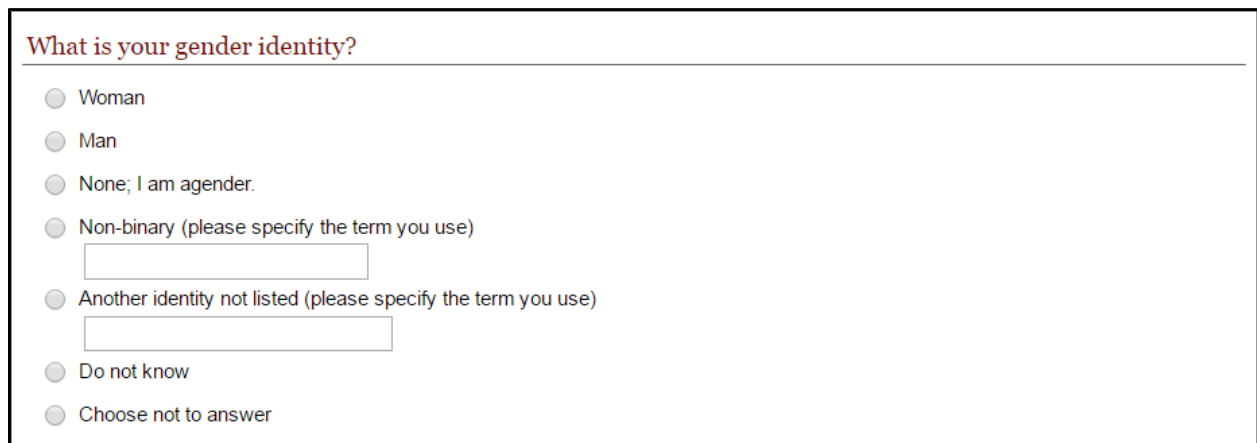
My adaptation of this question also incorporates the following changes from the version used by the Fenway Institute (2013):

- To simplify data analysis and maintain consistency across questions, this question is single-select (does not permit multiple selections).
- I omitted the separate options for “trans men” and “trans women” given that these respondents by definition identify as men or women, and because including a separate “transgender” option alongside “man” and “woman” suggests that researchers do not

perceive transgender people as men or women, producing identity threat for these respondents (Pope and Warner, 2016).

- Changing the labeling of the free-text box from “Additional Gender Category/Other” to “Another gender identity not listed” avoids literally “othering” participants who select that option. Instead, it places responsibility on the researcher for not including an option for that participant, and implicitly signals that the researcher is aware that other identities exist.
- I added a “Do not know” option to capture participants who are questioning their gender or who do not understand the question, following the recommendation of Redford and van Wagener (2012).

On the basis of participant feedback from Study 1, I slightly revised the question used in Study 2.



The image shows a screenshot of a survey question titled "What is your gender identity?". The question is presented in a red font. Below the title, there are seven radio button options: "Woman", "Man", "None; I am agender.", "Non-binary (please specify the term you use)", "Another identity not listed (please specify the term you use)", "Do not know", and "Choose not to answer". The "Non-binary" and "Another identity not listed" options each have a small text input box below them for specifying the term used.

Figure 4. Multiple-choice gender identity question used in Study 2.

A large number of participants who chose the “Another identity not listed” option in Study 1 wrote in “non-binary” or “agender,” so these were added as options in the version of this question used in Study 2. This new design also better matches the suggestion by Tate, Youssef, and Bettergarcia (2014) that gender identity can be summarized in relation to the dominant cultural categories of man and woman as man, woman, both, or neither. The revised version

replaces the term “genderqueer” with “non-binary,” which emerged as a commonly used umbrella term for a variety of gender identities that are not “man” or “woman.” To capture additional diversity within this umbrella category, I added an optional free-text box to the “non-binary” option.

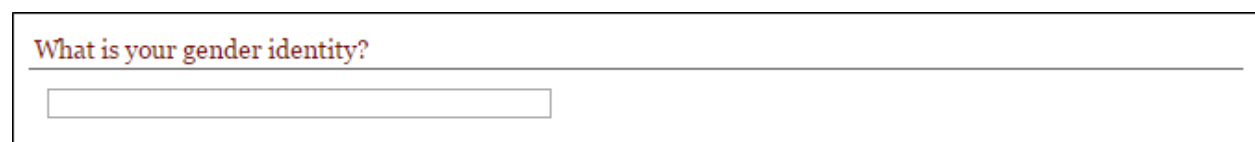
**Simplicity:** While it is slightly longer than the Binary + Other question, the Multiple-Choice question may be easier to analyze, as it includes more selection options to reduce the number of free-text responses that must be hand-coded.

**Validity:** In theory, this could be a valid measure of gender identity. It includes options for most major categories of gender identity (Tate et al., 2013).

**Acceptability:** This measure includes options for most common gender identity categories, signaling researchers’ awareness of and concern about these categories. This likely increases acceptability of this measure to gender minorities. However, it is possible that the inclusion of more than the two most common options may increase participant dissatisfaction if their identity is not specifically included. Cisgender participants may be confused by the number of additional options, although listing the familiar “Man” and “Woman” first may reduce this.

### **Free Text Box**

Many gender minority participants left comments in my pilot study requesting a free-text box in which to report their gender identities, so I included it for comparison in Studies 1 and 2.

The image shows a screenshot of a survey question. The question is "What is your gender identity?" displayed in a reddish-brown font. Below the question is a horizontal line, and underneath that line is a rectangular text input field with a thin border.

*Figure 5. Short free-text gender identity question used in Studies 1 and 2.*

This question gives participants complete control over what information to disclose about their genders. They are free to frame their identities exactly as they choose, and to disclose or

withhold any information at their discretion. This question makes no assumptions about gender identity, and thus appears perfectly inclusive. Data produced by this question require hand-coding before they are useable for statistical purposes, which places a substantial analytic burden on researchers. In addition, it requires researchers to determine how to group participants' responses rather than allowing them to group themselves. As the wording of the question does not give participants any information about the researchers' understandings of gender, participants are left to guess what level of disclosure will be appropriate. Thus, ironically, this apparently maximally inclusive question in practice gives participants very little agency over how their data are ultimately interpreted by researchers.

For further comparison, I also included a long version of the free-text question in Study 1. The paragraph-sized text box and the slight rewording of the prompt "Please describe your gender identity" normalizes the idea that gender identities are messy and require narratives rather than simple labels to understand.

The image shows a screenshot of a survey question. At the top, the text "Please describe your gender identity." is written in a reddish-brown font. Below this text is a large, empty rectangular text box with a thin grey border, intended for a long-form answer. The entire question is contained within a larger rectangular frame.

*Figure 6. Long free-text gender identity question used in Study 1.*

**Simplicity:** While free-text questions are visually simple, all responses require coding, making analysis of this question the most complex.

**Validity:** Like the Binary + Other question, this could theoretically be a valid measure of gender identity, because participants can write in any conceivable response. In practice, validity

will be determined more by the coding scheme developed by each researcher to organize the free-text responses than by the question itself.

**Acceptability:** As gender minority participants consistently request this specific measure, it seems to have very high acceptability. However, the size of the free text box and the slight rewording of the prompt relative to the short free text box question may make participants more aware that researchers are scrutinizing their genders, which may lead some participants to feel obligated to provide “proof,” a process which might induce feelings of identity threat. Cisgender people are unlikely to become confused by this question, as it does not include any unfamiliar terminology.

### **Measures of Gender Identity and Transgender Status**

Measures of gender identity alone cannot assess transgender status, as it is a separate construct. Therefore, a gender identity question must be combined with a second question measuring transgender status. For consistency, within each study, each measure used the same multiple-choice gender identity question paired with one of three different transgender status questions.

#### ***Multiple-Choice + Sex Assigned at Birth.***

This measure (i.e. a question about gender identity paired with a question about sex assigned at birth) was originally developed for public health and demographic research, and a growing body of evidence supports its use in that context (e.g. Sausa et al., 2009; Equality and Human Rights Commission, 2012; Redford & Van Wagenen, 2012; Tate et al., 2013; Fenway Institute, 2013; Conron et al., 2014). While the recommended wording for the gender identity options varies substantially by author, the sex assigned at birth question options are fairly consistent.

**What is your gender identity?**

Man

Woman

Genderqueer

Another identity not listed (please specify)

Do not know

Choose not to answer

---

**What was your sex assigned at birth?**

Male

Female

Intersex

Do not know

Choose not to answer

Figure 7. Multiple-choice gender identity question and sex assigned at birth question used in Study 1.

I used a slightly simplified version of the question wording recommended by the Fenway Institute (2013), and the options recommended by Tate et al. (2013), modified to add “Do not know” and “Choose not to Answer” options as recommended by Redford and Van Wagenen (2012). I wished to include an option for Intersex people as done by Tate et al. (2013), as several participants in my pilot study commented on the absence of such an option, but the Fenway Institute question wording no longer made sense with those options as a child cannot be assigned Intersex on a birth certificate in most parts of the world. Based on participant feedback from Study 1, however, I made further revisions to the question options to better acknowledge the fact that until very recently babies could not be assigned intersex at birth in the United States (Scutti, 2017).

**What is your gender identity?**

---

Man  
 Woman  
 None; I am agender.  
 Non-binary (please specify the term you use)  
  
 Another identity not listed (please specify the term you use)  
  
 Do not know  
 Choose not to answer

**What was your sex assigned at birth?**

---

Female  
 Male  
 Female, but I am intersex  
 Male, but I am intersex  
 Do not know  
 Choose not to answer

Figure 8. Multiple-choice gender identity question and sex assigned at birth question used in Study 2.

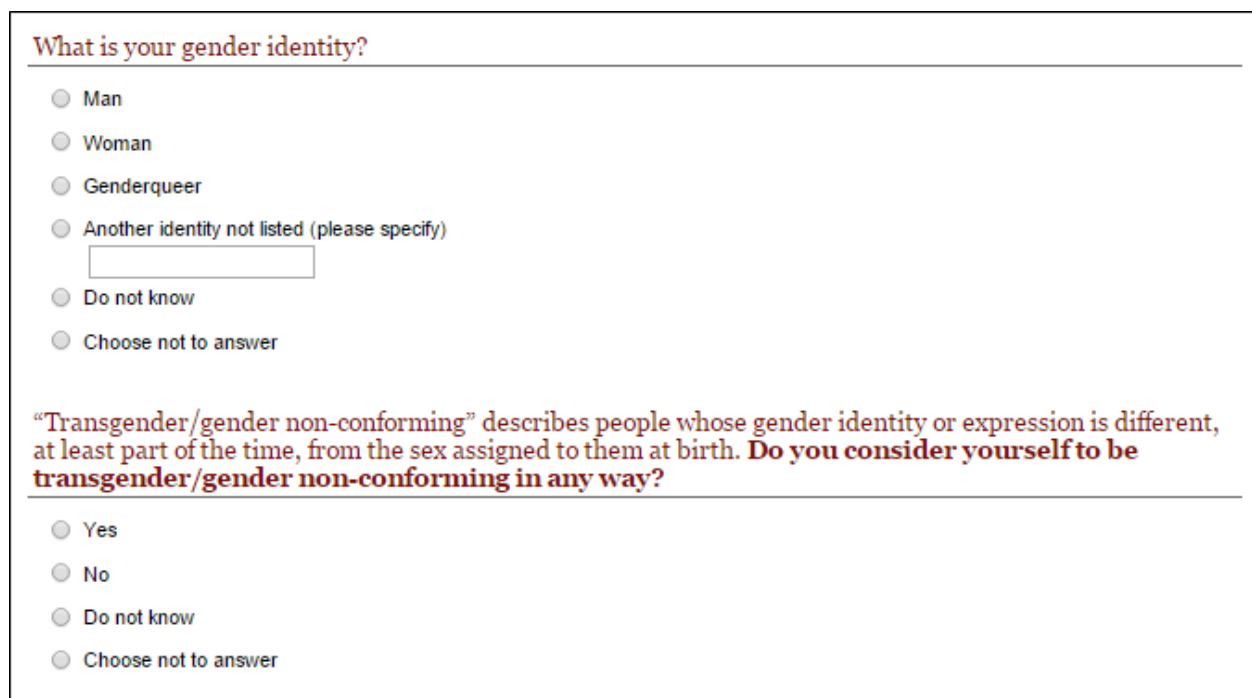
**Simplicity:** This measure is complex to analyze. People whose gender identity is the same as their sex assigned at birth are classified as cisgender, and people with other combinations of gender identity and sex assigned at birth are classified as transgender. Responses to the two questions must therefore be compared in a separate variable in order to classify participants as transgender or cisgender. In addition, it is unclear where to classify non-binary and agender participants (see Chapter 3).

**Validity:** Conron et al. (2014) suggest that some transgender participants may be misclassified as cisgender using this method, as they may report a sex assigned at birth that matches their gender identity, presumably to avoid the risk of being misgendered by researchers. Thus, the data produced by this question may not always be as valid as it appears.

**Acceptability:** As noted by Balarajan et al. (2011), cognitive interviewing suggests that some transgender people consider it offensive to be asked about their sex assigned at birth, even in health care contexts. This potential discomfort may be an acceptable trade-off in health care contexts, where a person’s body may be of as much interest as their identity and the need for invasive questions is clear to the participant. In the context of psychological research, however, the need for this information is not always clear, and other less-invasive questions may be equally useful and better-liked by participants.

### ***Multiple-Choice + Transgender Identity***

Many local and national transgender needs assessments have used this sort of question (e.g. Grant et al., 2011, Conron et al., 2012, Conron et al., 2014). For example, the National Transgender Discrimination Survey (Grant et al., 2011) used self-definition as a transgender person as their sole exclusion criterion, and I opted to use their question (with the addition of “do not know” and “choose not to answer” options) here.



The image shows a survey question with two parts. The first part asks "What is your gender identity?" and lists six radio button options: Man, Woman, Genderqueer, Another identity not listed (please specify) with a text input field below it, Do not know, and Choose not to answer. The second part provides a definition: "“Transgender/gender non-conforming” describes people whose gender identity or expression is different, at least part of the time, from the sex assigned to them at birth. Do you consider yourself to be transgender/gender non-conforming in any way?" and lists four radio button options: Yes, No, Do not know, and Choose not to answer.

*Figure 9. Multiple-choice gender identity question and transgender identity question used in Study 1.*



Based on participant feedback in Study 1, I removed the term “gender non-conforming” from the question stem, in order to ask a slightly narrower question about transgender identity in specific.

The image shows a survey question titled "What is your gender identity?". It features a list of radio button options: "Woman", "Man", "None; I am agender.", "Non-binary (please specify the term you use)" with a text input field, "Another identity not listed (please specify the term you use)" with a text input field, "Do not know", and "Choose not to answer". Below this is a second question: "\"Transgender\" describes people whose gender identity or expression is different, at least part of the time, from the sex assigned to them at birth. Do you consider yourself to be transgender?". This question also has radio button options for "No", "Yes", "Do not know", and "Choose not to answer".

Figure 10. Multiple-choice gender identity question and transgender identity question used in Study 2.

**Simplicity:** This measure is easier to analyze than the sex assigned at birth question, as transgender status can be determined with reference to a single question.

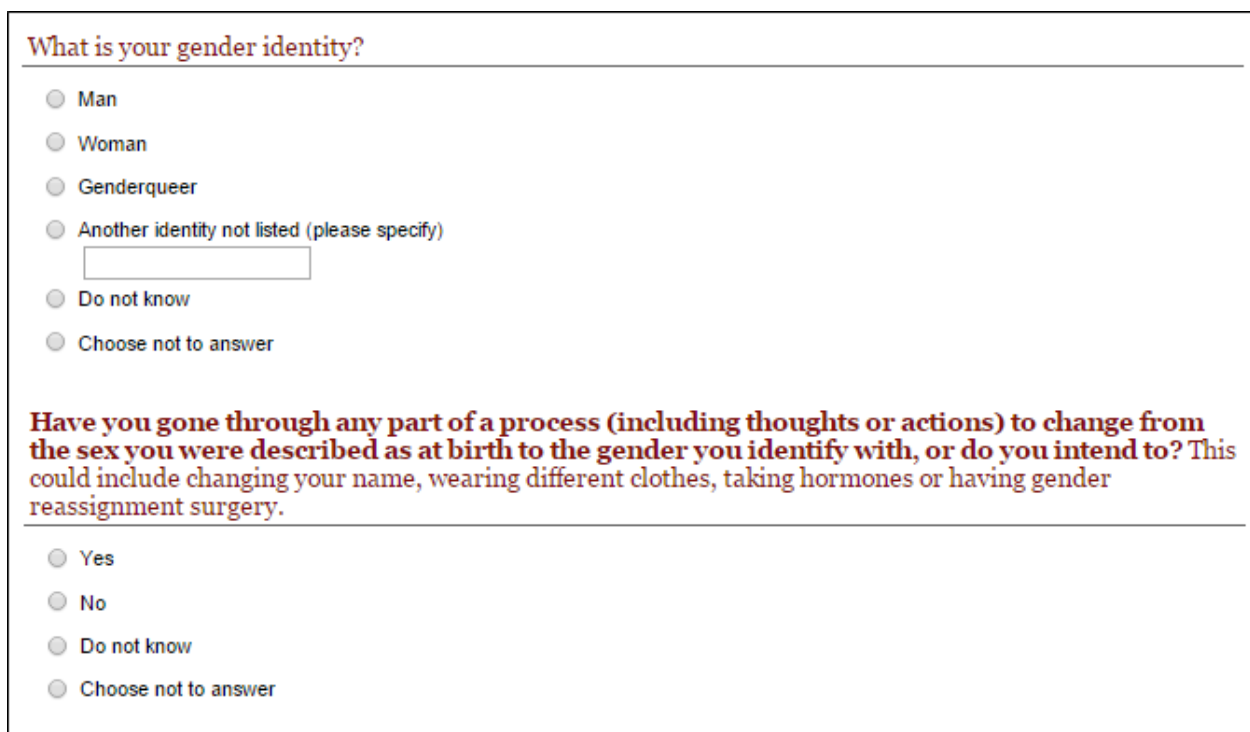
**Validity:** This question may overlook people who are not openly transgender or those who have a history of gender transition but do not identify as transgender (e.g. Tate et al., 2013). In contexts where a participant’s physical body is of interest (e.g. health care settings), this question may not gather the necessary information.

**Acceptability:** Unlike the Sex Assigned at Birth question, this question emphasizes self-identification over designations made by others, which may increase gender minority

participants' comfort and willingness to answer. However, cisgender people may be made uncomfortable or confused by being asked to consider whether they are transgender.

### ***Multiple-Choice + Transition History***

This measure assesses transgender status by asking about transition history. This method was developed by Balarajan et al. (2011), who suggest that it identifies a distinct subset of participants whom researchers may consider transgender but who are unlikely to be identified as such by other question types.



The image shows a screenshot of a survey question. The question is titled "What is your gender identity?" and has six radio button options: "Man", "Woman", "Genderqueer", "Another identity not listed (please specify)" (with a text input field below it), "Do not know", and "Choose not to answer". Below this is a second question: "Have you gone through any part of a process (including thoughts or actions) to change from the sex you were described as at birth to the gender you identify with, or do you intend to? This could include changing your name, wearing different clothes, taking hormones or having gender reassignment surgery." This question has four radio button options: "Yes", "No", "Do not know", and "Choose not to answer".

Figure 11. Multiple-choice gender identity question and transition history question used in Study 1.

**Simplicity:** Like the transgender identity measure, this measure is relatively simple to analyze because transgender status is determined by responses to single question.

**Validity:** This question can identify participants who have gone through gender transition but no longer identify as transgender. For some purposes, it may also identify a broader selection

of people as transgender than is desired by researchers, because its definition includes even thoughts about transition as sufficient to identify a person as transgender.

**Acceptability:** This question may be perceived as unnecessarily invasive in some contexts, as transition history is likely to be very private information for many gender minorities (Fenway Institute, 2013).

### **Conclusion**

In this chapter, I have reviewed previous attempts to develop inclusive measures of gender identity and transgender status, and identified several gaps in the evidence for the validity and acceptability of these existing measures. Based on the best available evidence, I then identified a few of the most promising measures which appear suitable for use in with both cisgender and transgender people and with people who have non-binary gender identities. For each of the measures, I identified a “best-case scenario” wording based on the insights of other researchers and comments from participants in several iterations of empirical testing. As no one measure can perfectly address all needs, I also briefly noted how simple each question is to analyze, how face valid it is as a measure of gender identity/transgender status, and how acceptable it is to gender minority and cisgender participants. In the next three chapters, I describe the results of two studies which present novel experimental evidence for the relative validity and acceptability of each of these questions.

### **Chapter 3: Simplicity of Inclusive Gender Identity and Transgender Status Measures**

I have just argued that researchers in social and health sciences should adapt their work to incorporate the now well-established idea that gender is multidimensional, and described previous attempts to create survey measures of gender identity and transgender status that meet this need. This chapter and the two that follow present two studies I conducted to identify the trans-inclusive gender question that best balances needs for *simplicity of administration and analysis* (Chapter 3), *acceptability to participants* (Chapter 4), and *validity* (Chapter 5).

To maximize the chances that a new measure will be adopted, it must be short and easy to administer and interpret. Binary gender questions are among the simplest possible survey questions to administer and interpret, so alternative measures will necessarily be more complex. However, complexity exists on a spectrum; questions which require extensive hand-coding (e.g. free-text questions) are more complex to analyze than those which require only programmatic comparison between two different questions to yield the needed information (e.g. the multiple-choice gender identity question and sex assigned at birth). By detailing the process I used to clean and organize the data from these two studies (including alternatives that I attempted but abandoned), I attempt to showcase the relative difficulty of using each method while also providing usage guidelines for future researchers who use these measures.

I also address in this chapter two components of validity; the need for a question to be *content valid* and *generally usable*. A measure which lacks content validity misses crucial information about the construct it is measuring, making the measure less sensitive or inaccurate. As this is the core criticism leveled against binary gender questions, I address it first. Additionally, I examine generalizability in this chapter because measures which are only useable

in a narrow context or which are not reliable are harder for researchers to use and less likely to be adopted. A simple measure should be usable across many contexts and participants.

Overall, my goal in this chapter is to identify the simplest question format which is still inclusive of transgender and non-binary people and to clearly describe a plan for using the data which it produces.

### **Assessing a Measure's Content Validity**

A measure of gender identity which has content validity does not omit any important facets of the construct of gender identity, nor include any extraneous elements (John & Benet-Martinez, 2000). To determine whether this is true, we need a precise theoretical definition of the construct(s) being measured. **I define *gender identity* as the relationship a person perceives between the self and the gender groups commonly recognized within their culture (see Chapter 1 for more detail).** I further propose that gender identity is the same construct for transgender and cisgender people, as well as for people with binary gender identities and non-binary gender identities (e.g. Tate, Youssef, & Bettergarcia, 2014); Joel et al., 2014; Olson et al., 2015).

A person's gender identity is not necessarily the same as external observers' perceptions of their gender group membership (e.g. Tate, Youssef, & Bettergarcia, 2014); rather, like other social identities, it is quintessentially based in self-categorization, a subjective sense of membership or lack of membership in a given gender group (e.g. Turner, 1982). According to social identity theory (Tajfel & Turner, 1979), self-categorization in a group results from assignment to that group. As established previously, due to definitional discrepancy between the gender group to which transgender people are assigned (their sex at birth) and their self-categorization, this view of the causes of self-categorization cannot account for or shed light on

the transgender experience. Therefore, I have moved away from social identity theory's conflation of self-categorization with assignment to a group, instead explicitly focusing on self-categorization as subjective membership.

Evidence for content validity can come partially from content experts (e.g. researchers with expertise in gender and transgender studies), who have designed measures to correspond to theoretical definitions of gender identity and transgender status. As the measures being tested here were already designed by such content experts, and often have gone through multiple iterations of such review, there is already some evidence for their content validity. To provide further evidence, I also chose to include in my studies two additional gender self-categorization measures designed to answer theoretical questions about the nature of binary and non-binary gender identities, and about the potential for overlapping gender identities.

Tate, Youssef, and Bettergarcia (2014) suggest that gender identity, in the sense of self-categorization, can be viewed as the extent of a person's identification with each of the available gender categories in their culture. Because gender in the United States is still largely understood as binary, the gender identity of a person in this culture can then be understood as the degree to which they perceive their identity as overlapping the category "men" and the category "women." Tate, Youssef, and Bettergarcia (2014) suggest that non-binary gender identities may either *integrate* both male and female identity (which they call genderblended) or *reject/transcend* male and female (which they call postgender). To assess this theory, in Study 1, I measured self-categorization among gender minority participants with a single item adapted from Joel et al. (2014), which reads as follows: *Currently, I think of myself as: (A man/A woman/Both a man and a woman/Neither a man nor a woman).*

In Study 2, in addition to the Joel et al. (2014) item, I measured self-categorization with a novel measure adapted from a racial identity measure developed by Rockquemore and Arend (2002). This measure was designed to capture the ways that the social interpretation and availability of certain identity labels may differ from person to person, as they are constrained by both personal characteristics of the individual (e.g. physical appearance) and the social environment (e.g. the available socially recognized identity categories). A combination of these factors seems to determine whether people with one black and one white parent identify themselves as black, white, and/or biracial, or indicate that they identify as black, white, or biracial, but experience the world as a black, white, or biracial person (Rockquemore, 1998). Likewise, a combination of personal and social characteristics may determine how gender minorities categorize their genders (e.g. as men, women, non-binary, agender, or something else).

Another source of content validity evidence is unconstrained responses from participants themselves. Typically, these are elicited through cognitive interviews. These yield information about how participants interpret the questions and choose their answers (e.g. Schwartz, 1999) to ensure that researchers and participants interpret the questions the same way, and that participants are only reporting information which is theoretically related to the construct being measured. While cognitive interviews have already been conducted for several of these measures, I wanted to provide similar evidence about the full range of participants' self-categorizations within a survey design, so I included a free-text question. If the majority of responses to the free-text question are captured by the options available in the multiple choice question, we have greater evidence of content validity (given my assumption that the free-text measure has the highest content validity from a participant's perspective).

## Assessing a Measure's General Usability

Demonstrating a measure's general usability, or the degree to which it can be used with other participants, tasks, and contexts, requires several different types of evidence. In this chapter, I consider the evidence for *reliability*. For instance, a measure must be reliable in order to be generally useable, so it must demonstrate adequate test-retest reliability, equivalence, and internal consistency. The question of the measure's general usability across participants and contexts is considered in Chapter 5.

*Test-retest reliability* (the relationship between the same measure given to the same participants at a different time) can in theory be easily demonstrated through mass testing and similar techniques. However, for the particular context of gender identity, any difference between the first and second tests could reflect changes in gender identity over time rather than a lack of reliability. Given the possibility that gender identity is a dynamically emerging construct (Diamond et al., 2011; Fausto-Sterling, 2012), approaches to establish test-retest reliability for a gender identity measure must be carefully considered. In both Study 1 and Study 2, I opted to administer the same gender identity measure to each person multiple times within the same survey, which provides some limited evidence of test-retest reliability. Alternatively, researchers may wish to simply accept that gender identity may differ from time to time and place to place in response to a wide variety of inputs (and may indeed even be influenced by the mere act of measuring it).

Evidence of *equivalence* (the relationship between different forms of the measure given to the same participants at the same time) might include the similarity between participants' scores on the various forms of inclusive questions. Both of these studies include multiple versions of the gender identity question with subtle differences, and while I do expect some



variation in responses to these questions (see Chapter 5), the vast majority of responses should remain consistent.

Finally, *internal consistency* is another aspect of reliability which is usually assessed using a split-halves or interrater technique. However, these methods (and the concept of internal consistency itself) are not particularly applicable to the proposed measure, as none of the items are presumed to be measuring exactly the same construct, and the measure is very short.

However, we will use independent raters to code the free-text portions of participants' responses, and can assess interrater reliability on that basis.

### **Summary of Goals**

The next portion of this chapter describes part of the methods for two studies conducted in 2016 with samples of undergraduate students and gender minorities to assess the simplicity, acceptability, and validity of several gender identity and transgender status measures in order to recommend a measure for general use in psychological research. In the Methods section, I present information about each sample and the gender identity and transgender status measures used in each study (see Chapter 2 for details about how these measures were selected). For each measure, I present the coding and/or analytic strategies I used for each one, with a specific focus on the handling of participants whose responses were challenging to categorize in a useable way. This section provides evidence about test-retest reliability, internal consistency, and inter-rater reliability. Then, in the Results section, I compare the categorizations produced by each measure, to determine how much of the information captured by each one is captured (or missed) by other measures, providing some evidence of the content validity of each one. Finally, in the Discussion section, I make general observations about the use of each method (including alternatives I did

not attempt that could be used by future researchers), and give specific recommendations about which measures are simplest to use.

## **Method**

### **Participants**

#### **Study 1 (Spring 2016).**

In spring 2016, I recruited a convenience sample of self-identified gender minorities<sup>9</sup> via social media, LGBT listservs, and communication from previous participants, and a separate convenience sample of undergraduate participants (unselected for gender identity or transgender status) from the psychology department participant pool at a large public university in the Midwest United States. Gender minority participants had the opportunity to enter a raffle to win one of 50 \$20 Amazon gift cards as compensation for participation (funding provided by a Clara Mayo Grant from the Society for the Scientific Study of Social Issues), while undergraduate participants received course credit for participating. Among the gender minority sample, 637 participants consented, 330 completed Section 1 of the survey, and 291 completed the entire survey. Thus, 46% of gender minority participants who viewed the survey link went on to complete the entire survey. This high dropout rate among the gender minority sample may be partially explained by the survey's length; it included well over 200 questions and likely took participants nearly an hour to complete. In the student sample, 191 participants consented and 190 completed the entire survey.

---

<sup>9</sup> The recruitment materials read as follows: "Help us improve demographic questions about gender! Do you identify as transsexual, transgender, genderqueer, non-binary, gender expansive, or gender non-conforming in some way, or have a history of gender transition? We want to hear your reactions to different types of gender questions to make these questions more inclusive and affirming for people with a variety of different gender identities and expressions. The study will take about 45 minutes. If you participate in this study, you can choose to be entered in a raffle to win one of fifty \$20 gift cards."

Gender minority participants (N = 256) were 5% Black (N = 13), 0.3% Pacific Islander (N = 1), 3% Middle Eastern (N = 7), 1% East Asian (N = 3), 1% Southeast Asian (N = 3), 89% White (N = 229), 3% Latinx (N = 8), and 1% Native American (N = 3). No participants reported Indian Subcontinent. In all samples across all studies, participants could select more than one race/ethnicity, so percentages sum to more than 100%. Sixteen percent (N = 40) selected more than one race, and 9% (N = 24) wrote in another racial identity (examples include Dutch, Jewish, Romani, and Singaporean Indian). Eight-seven percent (N = 225) had completed at least some college, 55% (N = 141) had a B.A. or higher, and 24% (N = 62) had a post-college degree. Sixty-six percent (N = 168) were lower middle class or below. Their ages ranged from 18-69 (M = 32).

Undergraduate participants (N = 190) were 8% Black (N = 15), 4% Indian Subcontinent (N = 7), 1% Middle Eastern (N = 2), 9% East Asian (N = 17), 1% Southeast Asian (N = 2), 80% White (N = 151), 5% Latinx (N = 10), and 2% Native American (N = 2). No participants reported Pacific Islander. Eleven percent (N = 20) selected more than one race, and 3% (N = 5) wrote in another racial identity (responses were Ethnically Persian/Iranian, Mexican, Prefer not to Answer, and White North American). Thirty percent were lower middle class or below. Despite the fact that all participants in this sample had to be enrolled as undergraduates in order to be part of the participant pool from which they were recruited, only 64% (N = 121) reported that they had completed at least some college, and 7% (N = 14) reported having a B.A. or higher. Their ages ranged from 18-22 (M = 19).

### **Study 2 (Fall 2016).**

In the fall of 2016, we again recruited a convenience sample of self-identified gender minorities<sup>10</sup> via social media (especially the Kinsey Institute Twitter and Facebook pages),

---

<sup>10</sup> The recruitment materials read as follows: “Help us improve demographic questions about gender! Do you identify as transsexual, transgender, genderqueer, non-binary, gender expansive, or gender non-conforming in some

LGBT listservs at major universities, and communication from previous participants, alongside a convenience sample of undergraduate participants (unselected for gender identity or transgender status) from the psychology department participant pool at a large public university in the Midwest United States. Gender minority participants had the opportunity to enter a raffle to win one of 50 \$20 Amazon gift cards as compensation for participation, and all undergraduate participants received course credit as compensation for participation. In the gender minority sample, 186 participants consented, 102 completed Section 1, and 83 completed the entire survey. Thus, 45% of gender minority participants who viewed the survey link went on to complete the entire survey. Again, this high dropout rate among the gender minority sample may be partially explained by the survey's length; Study 2 also included well over 200 questions and likely took participants nearly an hour to complete. In the student sample, 543 consented and 507 completed the entire survey, for a 93% completion rate.

Gender minority participants (N=82) were 4.8% Black (N=4), 1.2% Middle Eastern (N=1), 1.2% East Asian (N=1), 86.6% White (N=71), 4.8% Latinx (N=4), and 4.8% Native American (N=4).<sup>11</sup> Eight participants (9.8%) selected more than one race, and six (7.3%) wrote in another racial identity (responses were Ashkenazi, Central Asian, Latin American, mixed race – Black African/white, mixed white and Caribbean, and Slavic, Jew). Ninety-five percent had completed at least some college, 61% had a B.A. or higher, and 26% had completed a post-graduate degree. Their ages ranged from 18-74 (M = 29.8).

---

way, or have a history of gender transition? We want to hear your reactions to different types of gender questions to make these questions more inclusive and affirming for people with a variety of different gender identities and expressions. The study will take about 45 minutes. If you participate in this study, you can choose to be entered in a raffle to win one of fifty \$20 gift cards.”

<sup>11</sup> Participants could select more than one race/ethnicity, so percentages sum to more than 100%.

Undergraduate participants (N=505) were 8.3% Black (N=42), 0.3% Pacific Islander (N=2), 2.9% Indian Subcontinent (N=15), 1.7% Middle Eastern (N=9), 11.7% East Asian (N=59), 1.4% Southeast Asian (N=7), 73.3% White (N=370), 9.3% Latinx (N=47), and 0.8% Native American (N=4).<sup>12</sup> Forty-eight participants (9.5%) selected more than one race, and three (0.6%) wrote in another racial identity (responses were Jewish, Multicultural). As would be expected in a sample recruited from a student research pool, all undergraduate participants had completed at least some college, and 6% indicated that they had completed a B.A. or higher. Their ages ranged from 18-33 (M = 19).

## **Materials**

For a summary of all the measures included in Studies 1 and 2 and the order in which they appeared, refer to Appendix A.

In Study 1, participants completed five gender identity measures (Binary, Binary + Other, Multiple-Choice, Short Free Text, Long Free Text) and three gender identity and transgender status measures (Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Multiple-Choice + Transition History) for a total of eight different measures.

In Study 2, participants completed four gender identity measures (Binary, Binary + Other, Short Free Text, Multiple-Choice) and two gender identity and transgender status measures (Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity) for a total of six different measures. For more information about how these measures were selected and my predictions about each one, see Chapter 2.

---

<sup>12</sup> Participants could select more than one race/ethnicity, so percentages sum to more than 100%.

For comparison and to provide further evidence of content validity, participants also completed a measure of gender self-categorization from Joel et al. (2014) in Studies 1 and 2, and one adapted from Rockquemore and Arend (2002) in Study 2.

## Measures of Gender Identity.

### *Binary.*



What is your gender identity?

Man

Woman

Figure 1. Binary gender identity question used in Studies 1 and 2.

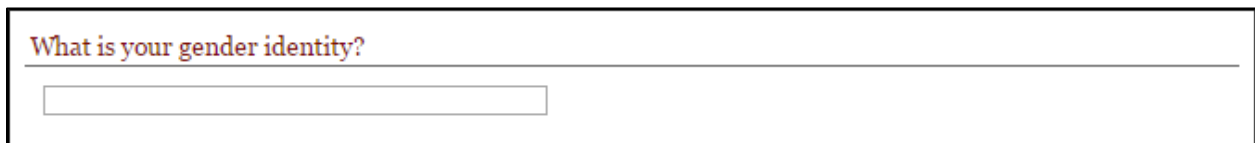
As described in Chapter 2, the binary question cannot yield complete information about gender identity or transgender status. It is included in these studies for comparison purposes.

Table 1

*Frequency of binary gender identities in Studies 1 and 2*

Sample	Gender Identity	
	Man	Woman
Study 1	203	264
Study 2	251	309

### *Short free text and long free text.*



What is your gender identity?

Figure 2. Short free-text gender identity question used in Studies 1 and 2.



Please describe your gender identity.

Figure 3. Long free-text gender identity question used in Study 1.

Responses to all free text gender questions were coded for gender identity and transgender status by three undergraduate coders (see Appendix A for the exact coding guidelines).<sup>13</sup> Coders grouped each free text response into one of six possible gender identity categories:

- **Man** (e.g. masculine, boy, transman, AMAB/DMAB, cis-male, cis-man).
- **Woman** (e.g. femme, girl, transwoman, AFAB/DFAB, cis-woman, cis-female).
- **Agender** (e.g. agender, none - preference is given to agender in descriptions such as “agender woman.”)
- **Non-binary** (e.g. androgynous, bigender, demiboy/demigirl, non-binary, genderfluid, genderqueer, genderfucker, neutrois, queer - preference is given to non-binary identities in descriptions such as “genderqueer woman” or “agender non-binary.”)
- **Rejects Gender** (e.g. gender is made up, no thanks, irrelevant).
- **Does not know** (e.g. I don’t know, not sure)
- **Not Enough Info** (e.g. human, weariness, cyborg, me, gendermeh, gender non-conforming).

For the short free-text question in Study 1, the three raters had 84% agreement for the gender minority sample (N = 282, Fleiss’ Kappa = .83, Z = 42.6, p <.001) and 98.9% agreement for the student sample (N = 186, Fleiss’ Kappa = .986, Z = 24.2, p <.001).

For the long free-text question in Study 1, the three raters had 71.3% agreement for the gender minority sample (N = 275, Fleiss’ Kappa = .69, Z = 36.4, p <.001) and 97.9% agreement for the student sample (N = 188, Fleiss’ Kappa = .973, Z = 24.9, p <.001). For the short free-text question in Study 2, the three raters had 85.5% agreement for the gender minority sample (N=83,

---

<sup>13</sup> We also coded the short free-text responses for transgender status. For Study 1, responses were classified as cisgender (N = 7), transgender (N = 52), or not enough information (N = 401), and of 460 free-text responses, 87% could not be classified by transgender status, presumably because the question did not explicitly request transgender status. For Study 2, responses were classified as cisgender (N = 303), transgender (N = 26), or not enough information (N = 209). Of the 538 total responses, 38% could not be classified by transgender status, presumably because the question did not explicitly request transgender status. As expected from the number of participants who were Unclassified by the Free Text question, there was relatively low agreement between the Free Text question and the other two questions (62% agreement between Transgender Identity and Free Text, and 67% agreement between Assigned Sex at Birth and Free Text). Taken together, these results suggest that the free text measure did not succeed at collecting transgender status information, so we did not conduct any additional analyses using this variable.



Fleiss' Kappa = .847,  $Z = 23$ ,  $p < .001$ ) and 99.8% agreement for the student sample ( $N = 459$ , Fleiss' Kappa = .997,  $Z = 37.8$ ,  $p < .001$ ).

To enable as many participants as possible to be categorized by free text response, we ultimately adopted a less conservative standard such that participants were placed in a gender identity category if two out of the three coders agreed. Even so, obtaining agreement this high required expertise about gender identity issues on the part of the raters. A majority of the research assistants who performed this coding had extensive personal familiarity with gender minority groups and high levels of knowledge about current gender terminology. Without this level of specialized knowledge, the interrater reliability would likely have been much lower.

Table 2

*Frequency of free-text gender identities in Studies 1 and 2*

Sample	Gender Identity by Free-Text							
	Man	Woman	Agender	Non-Binary	Rejects Gender	P Doesn't Know	Not Enough Info	Raters Disagreed
Study 1	140	135	17	145	7	7	9	12
Study 2	221	265	9	41	1	1	2	8

***Binary+Other.***

**What is your gender identity?**

---

Man  
 Woman  
 Other

*Figure 4. Binary plus Other gender identity question used in Studies 1 and 2.*

Participants who selected Other for the Binary + Other question were able to write in a free-text response. These responses were coded the same way as the free-text questions, by three undergraduate coders (see Appendix for the exact coding guidelines).

For the free-text responses to the Binary + Other question in Study 1, the three raters had 88.8% agreement for the gender minority sample (N = 187, Fleiss' Kappa = .83, Z = 33.9, p <.001). In the student sample, the raters had perfect agreement for the one participant who selected "other," and reliabilities could not be calculated. In Study 2, the three raters had 84.9% agreement for the gender minority sample (N = 53, Fleiss' Kappa = .77, Z = 16.8, p <.001) and 100% agreement for the student sample (N = 3, Fleiss' Kappa = 1.0, Z = 4.24, p <.001).

In Study 1, 12 participants who selected Other were coded as Man based on their free-text responses ("Transman," "transgender man," "transmasculine," "Intersex Man," "Female-bodied Man"), and another 11 were coded as Woman ("trans woman," "Transgender Woman," "Butch," "Butch Woman," "FTMTF"). In Study 2, two participants who selected Other were coded as Man based on their free-text responses ("trans masculine," "transgender, masculine"), and another two who selected Other were coded as Woman ("Transgender woman," "transwoman"). They were regrouped with the people who initially selected Man and Woman (but see the Discussion for a consideration of the issues with this recategorization).

Table 3

*Frequency of binary + other gender identities in Studies 1 and 2*

Sample	<b>Gender Identity by Binary + Other</b>							
	Man	Woman	Other <sup>a</sup>	Agender	Non-Binary	Rejects Gender	P Doesn't Know	Not Enough Info
Study 1	155	154	17	17	128	4	2	5
Study 2	233	284	6	7	39	-	3	3

<sup>a</sup>People in this category selected Other but did not write in anything for the associated free-text response.

***Multiple-choice.***

What is your gender identity?

---

Man  
 Woman  
 Genderqueer  
 Another identity not listed (please specify)  
  
 Do not know  
 Choose not to answer

Figure 5. Multiple-choice gender identity question used in Study 1.

What is your gender identity?

---

Woman  
 Man  
 None; I am agender.  
 Non-binary (please specify the term you use)  
  
 Another identity not listed (please specify the term you use)  
  
 Do not know  
 Choose not to answer

Figure 6. Multiple-choice gender identity question used in Study 2.

In Study 1, participants answered the same single-select multiple-choice question four times; in isolation, followed by a question about sex assigned at birth, followed by a question about transgender identity, and followed by a question about transition history. In Study 2, participants answered the same multiple-choice question three times; in isolation, followed by a question about sex assigned at birth, and followed by a question about transgender identity. We compared responses to each of the identical multiple-choice questions to classify participants by gender identity.

*Coding inconsistent responses.*

We initially took a conservative approach; only those who responded the same way to all of the questions (all four questions in Study 1, all three questions in Study 2) were categorized as

that gender identity for the subsequent analyses, while those who responded inconsistently across the questions (including those who skipped one of the multiple-choice questions) were classified as “Inconsistent.” Using this analytic strategy, 9% of the participants in Study 1 (N=42) and 11% of the participants in Study 2 (N=67) were classified as “Inconsistent,” usually because they skipped one question.

We felt that this definition was too conservative, so we revised our approach to classify anyone who responded consistently to three out of four multiple-choice questions (in Study 1) or to two out of three multiple-choice questions (in Study 2) as that gender identity. Participants who left three questions blank (in Study 1) or two questions blank (in Study 2) were also classified as inconsistent. Under this new definition, only 4% of participants in Study 1 (N=17) and 3% of participants in Study 2 (N=19) responded inconsistently; these responses are shown in Tables 4 and 5. For Study 1, 94% of the inconsistencies (N=16) were in the gender minority sample, but for Study 2, only 47% (N=9) were.

Table 4

*Inconsistent responses to multiple-choice gender identity measure (Study 1).*

<b>Sample</b>	<b>Multiple-Choice</b>	<b>MC + Assigned Sex at Birth</b>	<b>MC + Transgender Identity</b>	<b>MC + Transition History</b>	<b>Free-Text</b>
	Other	Genderqueer	Woman	Woman	Sometimes I'm a lady and sometimes I have no gender.
	Man	-	-	-	
	Choose Not to Answer	Other	-	-	
	Choose Not to Answer	Choose Not to Answer	Other	Don't Know	
	Other	Man	Man	Other	Genderqueer trans male
	-	Man	-	-	
	Don't Know	Don't Know	Woman	Woman	Assigned female, masculine presenting most of the time, confused about my gender identity
Gender minority	Genderqueer	Genderqueer	-	-	
	Genderqueer	Genderqueer	Don't Know	Other	trans-fabulous bottom with a touch of raccoon
	Woman	-	-	-	
	Genderqueer	Genderqueer	Woman	Woman	Agender Trans woman
	Choose Not to Answer	Genderqueer	Genderqueer	Choose Not to Answer	Genderfluid, typically leaning male
	Woman	Woman	-	-	
	Choose Not to Answer	Don't Know	Choose Not to Answer	Don't Know	
	-	-	Choose Not to Answer	-	
	Other	Don't Know	Other	Don't Know	jfc you tell me i dont know man i dont have one
Student	Man	-	Man	-	male

Table 5

*Inconsistent responses to multiple-choice gender identity measure (Study 2)*

<b>Sample</b>	Multiple-Choice	MC + Assigned Sex at Birth	MC + Transgender Identity	Free-Text
Gender minority	Don't know	-	Woman	bisexual
	Other	Non-binary	Man	I hate gender
	-	-	Other	I don't have one, I don't feel like gender should matter (though yes it does). I am somewhere in the spectrum of genders that is possible.
	-	Don't know	-	Non-binary woman
	-	Woman	-	-
	-	-	Woman	-
	-	Man	-	Trans man
	Non-binary	-	-	agender/male
	-	-	Non-binary	male-brained in a female body, but not gender dysphoric
	-	Man	-	-
Student	Woman	-	-	Woman
	Man	-	-	Male
	-	-	Woman	-
	-	-	Man	Male
	-	Woman	-	-
	-	Man	-	male
	-	Woman	-	-
	Man	-	-	Male
	-	Man	-	-

We examined the free-text responses of participants whose gender identity was classified as “Inconsistent” to determine whether their inconsistent responding reflected genuine uncertainty about how to classify their gender identity on the multiple-choice question, or was merely an artifact of completing the survey quickly and without much thought. We found distinctly different patterns in the gender minority and student samples. As shown in Tables 4 and 5, many responses from the gender minority sample (e.g. “Assigned female, masculine presenting most of the time, confused about my gender identity”) seem to indicate a substantial level of uncertainty on the part of the participant about how best to represent their gender identities using a multiple-choice question (an uncertainty that in many cases was also present for our coders when attempting to classify their free-text responses). In contrast, free-text responses of the “inconsistent” people in the student sample included “female,” “male,” “heterosexual woman,” and “woman?,” which seems to suggest that in this sample inconsistency was more often caused by hasty responding or the assumption that the researchers already had this information than by actual uncertainty about gender identity.

Even among participants who were classified as Inconsistent, very few reported actually differing gender identities on across questions. In Study 1, seven participants were inconsistent because they left two or more of the questions blank, and two more were inconsistent because they selected “Choose Not to Answer” for two or more of the questions. For Study 2, 17 of the 19 inconsistent participants were classified that way because they left two or more questions blank. This pattern of responding seems likely to be an artifact of survey fatigue, which we would not expect to influence participant responses to a gender question under ordinary circumstances. An alternative explanation is that participants felt more willing to answer some questions than others, or were using non-response to express dissatisfaction with some aspect of

the question and/or the survey overall; for a detailed exploration of this possibility, refer to Chapter 4.

*Coding “Another identity not listed” responses.*

For participants who consistently responded “Another identity not listed” (Study 1 N = 105; Study 2 N = 10) we compared the codes assigned to the free text response they entered and added those participants into one of the existing categories where possible to facilitate their inclusion in statistical analyses (for interrater reliabilities, see the Appendix). This was particularly relevant in Study 1, where the multiple-choice gender identity question did not include an option for non-binary (although genderqueer was an option). Many participants wrote in identities like “non-binary” or “genderfluid,” and we grouped them with those who selected “genderqueer,” while participants who wrote in “transgender woman” or “heterosexual woman” were combined with those who selected “woman.”

In Study 1, 14 participants also wrote in “agender,” which seemed conceptually distinct from non-binary (see Tate, Youssef, & Bettergarcia, 2014 for rationale) and so “agender” was added as a separate category (and included as its own option in Study 2). Participants for whom there was not enough information to make a categorization (e.g. “????????????????”) and those whose responses indicated a rejection of gender (e.g. “I do not identify with gender roles at all”) were left in the “Another identity not listed” category. Participants whose responses could not be categorized in any of the previous ways (i.e. those who selected “Another identity not listed” and were not categorized the same way on three out of four of their free-text responses to the multiple-choice gender questions) were grouped with those who inconsistently answered the multiple-choice questions.



Participants were ultimately grouped into one of eight possible categories on the basis of their multiple-choice gender identity (see Table 6). To allow adequate cell sizes for certain statistical analyses, we also created a condensed version of multiple-choice gender identity that dropped all categories except Man, Woman, and Non-Binary/Agender.

Table 6

*Frequency of multiple-choice gender identities in Studies 1 and 2*

<b>Gender Identity by Multiple-Choice</b>								
<b>Sample</b>	Man	Woman	Agender	Non-Binary	Another Identity Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Study 1	144	142	14	139	7	11	7	20
Study 2	243	282	8	47	-	2	5	19

**Alternative Measures of Gender Self-Categorization.**

As a pilot test, the gender minority sample in Study 1 completed the following single-select self-categorization measure developed by Joel et al. (2014): *Currently, I think of myself as: A man/A woman/Both a man and a woman/Neither a man nor a woman.* All participants completed it in Study 2.

Table 7

*Frequency of Gender Self-Categorization (Joel et al., 2014) in Study 1 Gender Minority Sample and Study 2.*

Sample	Gender Self-Categorization (Joel et al. 2014)			
	Man	Woman	Both	Neither
Study 1	45	64	43	120
Study 2	248	283	19	38

In Study 2, participants also completed the following single-select multiple choice question (adapted from Rockquemore and Arend, 2002):

*Please indicate which of the following choices best describes how you personally think of your gender, even if it is not always how you are perceived by others.*

- *I consider myself exclusively a woman*
- *I consider myself exclusively a man.*
- *I consider myself exclusively non-binary (neither a man nor a woman).*
- *I consider myself non-binary, but I experience the world as a woman.*
- *I consider myself non-binary, but I experience the world as a man.*
- *I sometimes consider myself a man or a woman, and sometimes non-binary depending on the circumstances.*
- *Gender is meaningless. I do not believe in gender identities.*
- *Something else (please describe). [free text box]*

Table 8

*Frequency of Gender Self-Categorization (Rockquemore & Arend, 2002) in Study 2*

Gender Self-Categorization (Rockquemore & Arend, 2002)							
Man <sup>a</sup>	Woman <sup>b</sup>	Binary/NB <sup>c</sup>	NB, Man <sup>d</sup>	NB, Woman <sup>e</sup>	NB <sup>f</sup>	Rejects Gender <sup>g</sup>	Something Else <sup>h</sup>
232	267	11	5	30	12	14	21

## Measures of Gender Identity and Transgender Status.

### *Multiple-Choice + Sex Assigned at Birth.*

What is your gender identity?

- Man
- Woman
- Genderqueer
- Another identity not listed (please specify)
- Do not know
- Choose not to answer

What was your sex assigned at birth?

- Male
- Female
- Intersex
- Do not know
- Choose not to answer

Figure 7. Multiple-choice gender identity and sex assigned at birth question used in Study 1.

What is your gender identity?

- Man
- Woman
- None; I am agender.
- Non-binary (please specify the term you use)
- Another identity not listed (please specify the term you use)
- Do not know
- Choose not to answer

What was your sex assigned at birth?

- Female
- Male
- Female, but I am intersex
- Male, but I am intersex
- Do not know
- Choose not to answer

*Figure 8. Multiple-choice gender identity and sex assigned at birth question used in Study 2.*

We compared the multiple-choice gender identity category to which each participant was assigned (not just their response to this measure; see previous section) to the sex assigned at birth question in order to categorize participants as Not Transgender, Transgender, or Unclassified. Participants whose gender identity was Man and sex assigned at birth was Male, or whose gender identity was Woman and sex assigned at birth was Female, were categorized as Not Transgender. Participants whose gender identity was Man and sex assigned at birth was Female, or whose gender identity was Woman and sex assigned at birth was Male, were categorized as Transgender. Participants with any other combination of gender identity and sex assigned at birth were categorized as Unclassified.<sup>14</sup>

*Coding transgender status for non-binary and agender participants.*

Importantly, in this categorization scheme, non-binary and agender people appear in the Unclassified category, a significant change from typical practice. In a typical coding scheme, either all non-binary and agender people are transgender (if transgender is defined as having any gender identity different from sex assigned at birth), or no non-binary or agender people are transgender (if transgender is defined to include only people with binary gender identities). Such categorization structures make transgender status redundant with gender identity; that is, knowing a non-binary or agender person’s sex assigned at birth does not provide any additional information about their transgender status that was not already evident from their gender identity.

Table 9  
*Frequency of transgender status by gender identity + sex assigned at birth in Studies 1 and 2.*

Sample	Transgender Status by Gender Identity + Sex Assigned at Birth		
	Not Transgender	Transgender	Unclassified

<sup>14</sup> In Study 1, only three participants selected Intersex, and in Study 2 one participant selected “Female, but I am intersex” and one participant selected “Male, but I am Intersex” while no participants selected “Intersex,” supporting the revised wording. No participants in Study 1 and only one participant in Study 2 selected “Do not know.”

---

Study 1	214	57	33
Study 2	495	20	76

---

*Multiple-Choice + Transgender Identity.*

What is your gender identity?

Man

Woman

Genderqueer

Another identity not listed (please specify)

Do not know

Choose not to answer

---

“Transgender/gender non-conforming” describes people whose gender identity or expression is different, at least part of the time, from the sex assigned to them at birth. **Do you consider yourself to be transgender/gender non-conforming in any way?**

Yes

No

Do not know

Choose not to answer

Figure 9. Multiple-choice gender identity question and transgender identity question used in Study 1.

What is your gender identity?

Woman

Man

None; I am agender.

Non-binary (please specify the term you use)

Another identity not listed (please specify the term you use)

Do not know

Choose not to answer

---

"Transgender" describes people whose gender identity or expression is different, at least part of the time, from the sex assigned to them at birth. **Do you consider yourself to be transgender?**

No

Yes

Do not know

Choose not to answer

Figure 10. Multiple-choice gender identity question and transgender identity question used in Study 2.

We categorized participants into one of three transgender status groups based on their answer to the transgender identity question. Those who answered “No” are categorized as Not Transgender, those who answered “Yes” are categorized as Transgender, and those who indicated “Do not know” or “Choose not to answer” are combined into an Unclassified category. As the transgender identity question categorizes participants’ transgender status based exclusively on their identity, non-binary and agender people may appear in any category.

Table 10

*Frequency of transgender status by transgender identity in Studies 1 and 2.*

<b>Transgender Status by Transgender Identity</b>			
<b>Sample</b>	Not Transgender	Transgender	Unclassified
Study 1	255	196	15
Study 2	506	62	19

**Multiple-Choice + Transition History.**

What is your gender identity?

Man  
 Woman  
 Genderqueer  
 Another identity not listed (please specify)  
  
 Do not know  
 Choose not to answer

**Have you gone through any part of a process (including thoughts or actions) to change from the sex you were described as at birth to the gender you identify with, or do you intend to? This could include changing your name, wearing different clothes, taking hormones or having gender reassignment surgery.**

Yes  
 No  
 Do not know  
 Choose not to answer

Figure 11. Multiple-choice gender identity and transition history measure used in Study 1.

We categorized participants into one of three transgender status groups based on their answer to the transition history question. Those who answered “No” are not transgender, those who answered “Yes” are transgender, and those who indicated “Do not know” or “Choose not to answer” are combined into an Unclassified category. As the transition history question categorizes participants’ transgender status based exclusively on their thoughts and actions related to transitioning, non-binary and agender people may appear in any category.

Table 11

*Frequency of transgender status by transition history in Study 1*

<b>Transgender Status by Transition History</b>		
Not Transgender	Transgender	Unclassified
220	101	22

**Results**

**Comparison of Binary and Multiple-Choice Gender Identity Measures**



In both samples, the multiple-choice gender identity question suggests a much more gender-diverse sample than the binary question (Tables 12 and 13). In Study 1, 34.3% (N = 158 of 460) participants were consistently classified as a gender other than man or woman, and in Study 2, 7.3% (N = 41 of 557) were (bearing in mind that the student sample made up a much larger proportion of participants in Study 2 than Study 1).

Table 12

*Gender identity classification by binary and multiple-choice measures (Study 1; N = 460)*

<b>Binary</b>	<b>Multiple-Choice</b>							
	Man	Woman	Agender <sup>a</sup>	Gender-queer <sup>b</sup>	Another Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Man	<b>144</b>	-	4	35	2	6	1	10
Woman	-	<b>142</b>	9	88	3	4	6	6

<sup>a</sup>Agender was not included as an explicit category in the multiple-choice question in the Study 1 survey, but was added because so many participants wrote it in to the “Another identity not listed” category.

<sup>b</sup>The option in the multiple-choice question was Genderqueer, but people who selected Another Identity Not Listed and wrote in “non-binary” or similar also appear in this column.

Table 13

*Gender identity classification by binary and multiple-choice measures (Study 2; N = 557)*

<b>Binary</b>	<b>Multiple-Choice</b>							
	Man	Woman	Agender	Non-Binary	Another Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Man	<b>233</b>	-	2	7	-	-	4	5
Woman	-	<b>272</b>	3	22	-	2	1	6

### **Comparison of Binary + Other and Multiple-Choice Gender Identity Questions**

Table 14 shows how each participant’s gender identity was classified by the binary + other and multiple-choice gender identity measures for Study 1, and Table 15 shows the same information for Study 2. In Study 1, 85% (N = 409 of 480) were classified into the same one of our main four gender categories by both methods, and in Study 2, 95.4% (N = 547 of 573) were. These participants appear in green cells with bold text. A further 1% (N = 5) of participants in Study 1 and 0.1% (N=1) in Study 2 did not provide enough information, indicated that they did

not know their gender identity, or could not be classified on both measures; they appear in red cells with bold text.

### **Inconsistent Responses between Binary + Other and Multiple-Choice Measures.**

#### *Green Quadrant.*

The green quadrant shows responses that could be classified into one of our four main gender identity categories by both methods. The participants who were consistently classified have already been discussed.

In Study 1, 3.1% (N = 15) of participants who were classified as one of the four main genders by either method were classified as a different one of the main four genders by the other method. Two participants classified as men and seven classified as women based on the binary + other question (and three who wrote in “agender”) selected Genderqueer for the multiple-choice gender measure. Three more participants who were classified as non-binary from their responses to the binary + other question selected Man or Woman on the multiple-choice measure.

In Study 2, 1% (N = 6) of participants who were classified as one of the four main genders by either method were classified as a different one of the main four genders by the other method. One participant selected “Other” for the Binary + Other question and wrote in “agender,” and one other participant wrote in “transgender, masculin” and was recategorized as Man based on that response. Another participant selected “Man” for the binary + other question but “Genderqueer” for the multiple-choice questions and wrote in “male” for the short free-text measure. Finally, three participants selected “Woman” for the binary + other question, but “Man” for the multiple-choice questions, and wrote in “male” or “man” for the short free-text measure.

#### *Yellow Quadrant*

The yellow quadrant shows participants who were classified as one of the four main genders by the multiple choice questions but not by the binary + other question.

In Study 1, 2.7% of participants (N=13) were in this quadrant. All thirteen of them selected Other for the binary + other question, but did not write anything in the accompanying text box and so could not be classified any further by this question. Twelve of these participants selected Genderqueer for the multiple choice measure, and one wrote in agender.

In Study 2, 1% (N = 6) of participants were in this quadrant. All six of these participants selected Other for the binary + other question, but four of them did not write anything in the accompanying text box and so could not be classified any further by this question. Two of them selected “Man” and two selected “Non-Binary” for the multiple-choice question, and they wrote in “Man,” “Agender/possibly male,” “Non-binary, demiboi, transmasculine,” and “genderqueer” for the short free-text measure. The remaining two participants wrote “no” and “I’m not sure, transgender?” in the text box for the binary + other question, selected “Agender” and “Woman” for the multiple-choice question, and wrote in “agender” and “woman?” for the short free-text measure.

### *Blue Quadrant*

The blue quadrant shows participants who were classified as one of the four main gender identities by the binary + other question but not by the multiple choice question.

In Study 1, 5.8% of participants (N=28) were in this quadrant. For the binary + other question, eleven selected man and seven selected woman, nine were classified as non-binary based on their write-in response, and one was classified as agender. On the multiple-choice questions, fifteen were inconsistently classified (responses included “genderfluid,” “non-binary trans boy,” and “sometimes I’m a woman and sometimes I’m something else entirely”), five

chose not to answer, and seven indicated that they didn't know their identity (responses included "middleish").

In Study 2, 1.3% (N = 8) of participants were in this quadrant. For the binary + other question, three selected "Man" and five selected "Woman." For the multiple choice question, one selected Don't Know and two selected Choose Not to Answer; their short free-text responses were "Female," "Male," and "Female," respectively. The remaining five were inconsistently classified by the multiple choice question. Two wrote in "Male" for the short free-text response, one wrote in "Bisexual," and two did not write anything.

### *Red Quadrant*

The red quadrant shows participants who could not be classified as one of the four main gender identities by either the multiple-choice questions or the binary + other question.

In Study 1, 3.1% (N = 15) of participants were in this quadrant. All 15 selected other for the binary + other question, and four of them did not write in anything. Five did not give enough information to allow further classification (responses included "SirHer," "genderfuck," "transgender," "binary," and "transgender queer"). Two were classified as not knowing their gender identity (responses included "?????????????" and "don't know"), and four were classified as rejecting gender (responses included "Ethically and biologically irrelevant in nearly all contexts," "human," "gender does not exist" and "gender irrelevant").

In Study 2, 1% (N = 6) of participants were in this quadrant. All six selected Other for the binary + other question, and two of them did not write anything in the associated text box, so they could not be categorized further by that question. For the short free-text question, they wrote "trans masculine" and "agender/male." The remaining four wrote in "Questioning," "shruggie," "Man and Woman are not gender identities. They are biological sex," and "Who

gives a shit? Sorry” for the binary + other question. For the short free-text question, these participants wrote in “Questioning,” “I hate gender,” nothing (left the question blank), and “I don't have one, I don't feel like gender should matter (though yes it does). I am somewhere in the spectrum of genders that is possible.”

Table 14

*Gender identity classification by binary+other and multiple-choice measures (Study 1; N = 480)*

<b>Binary + Other</b>	<b>Multiple-Choice</b>							
	Man	Woman	Agender <sup>a</sup>	Gender-queer <sup>b</sup>	Another Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Man	<b>141</b>	-	-	2	-	4	1	6
Woman	-	<b>140</b>	-	7	-	-	3	4
Agender	-	-	<b>13</b>	3	-	-	-	1
Non-Binary	1	2	-	<b>115</b>	1	3	1	4
Other <sup>c</sup>	-	-	1	12	-	1	1	2
Not Enough Info	-	-	-	-	<b>3</b>	1	1	-
P Doesn't Know	-	-	-	-	-	<b>2</b>	-	-
Rejects Gender	-	-	-	-	3	-	-	1

*Note.* No participants were classified as Raters Disagreed based on their free-text responses to the binary + other question, so that row was omitted.

<sup>a</sup>Agender was not included as an explicit category in the multiple-choice question in Study 1, but was added because so many participants wrote it in to the “Another identity not listed” category.

<sup>b</sup>The option in the multiple-choice question was Genderqueer, but people who selected Another Identity Not Listed and wrote in “non-binary” or similar also appear in this column.

<sup>c</sup>People in this category selected Other but did not write in anything for the associated free-text response.

Table 15

*Gender identity classification by binary+other and multiple-choice measures (Study 2; N = 573)*

<b>Binary + Other</b>	<b>Multiple-Choice</b>							
	Man	Woman	Agender	Non- Binary	Another Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Man	<b>228</b>	-	-	2	-	-	1	2
Woman	3	<b>274</b>	-	-	-	1	1	3
Agender	-	-	<b>6</b>	1	-	-	-	-
Non-Binary	-	-	-	<b>39</b>	-	-	-	-
Other <sup>a</sup>	2	-	-	2	-	-	1	1
Not Enough Info	-	-	1	-	-	-	1	1
P Doesn't Know	-	1	-	-	-	<b>1</b>	-	1
Rejects Gender	-	-	-	-	-	-	-	-

*Note.* No participants were classified as Rejects Gender or Raters Disagreed based on their free-text responses to the binary + other question, so those rows were omitted.

<sup>a</sup>People in this category selected Other but did not write in anything for the associated free-text response.

## Comparison of Free-Text and Multiple-Choice Gender Identity Measures

Table 16 shows how each participant's gender identity was classified by the short free-text<sup>15</sup> and multiple-choice gender identity measures for Study 1, and Table 17 shows the same information for Study 2. In Study 1, 87.6% (N = 404 of 461) were classified into the same one of our main four gender categories by both methods, and in Study 2, 95.7% (N = 516 of 539) were. These participants appear in green cells with bold text. A further 1.7% (N=8) of participants in Study 1 and 0.4% (N=2) in Study 2 did not provide enough information, indicated that they did not know their gender identity, or could not be classified on both types of measures; they appear in red cells with bold text. Overall, the vast majority of participants were classified the same way by both the multiple-choice and free text gender identity question across both studies.

---

<sup>15</sup> The short and long free-text responses were usually coded the same way, so for the sake of space we focused only on the short free-text question (which seems more likely to be widely used in any case).



Table 16

*Gender identity classification by free-text and multiple-choice measures (Study 1)*

Free Text	Multiple-Choice							
	Man	Woman	Agender <sup>a</sup>	Gender-queer <sup>b</sup>	Another Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Man	<b>137</b>	-	-	-	-	1	1	1
Woman	-	<b>128</b>	-	5	1	-	1	-
Agender	-	-	<b>13</b>	2	-	1	-	1
Non-Binary	3	8	-	<b>126</b>	1	1	1	5
Not Enough Info	-	1	-	1	<b>3</b>	1	2	1
P Doesn't Know	-	-	-	-	-	<b>5</b>	-	2
Rejects Gender Raters Disagreed	-	2	-	-	2	1	1	1
	-	-	-	-	1	-	-	-

<sup>a</sup>Agender was not included as an explicit category in the multiple-choice question in Study 1, but was added because so many participants wrote it in to the “Another identity not listed” category.

<sup>b</sup>The option in the multiple-choice question was Genderqueer, but people who selected Another Identity Not Listed and wrote in “non-binary” or similar also appear in this column.

Table 17

*Gender identity classification by free-text and multiple-choice measures (Study 2)*

Free Text	Multiple-Choice							
	Man	Woman	Agender	Non-Binary	Another Not Listed	Don't Know	Choose Not to Answer	Inconsistent
Man	<b>212</b>	-	-	1	-	-	2	6
Woman	-	<b>259</b>	-	1	-	1	1	1
Agender	1	-	<b>6</b>	1	-	-	-	1
Non-Binary	-	1	-	<b>39</b>	-	-	-	1
Not Enough Info	-	-	-	-	-	-	-	1
P Doesn't Know	-	-	-	-	-	<b>1</b>	-	-
Rejects Gender	-	-	-	-	-	-	-	1
Raters Disagreed	-	-	-	-	-	-	1	<b>1</b>

## **Inconsistent Responses Between Multiple-Choice and Free-Text Measures.**

### *Green Quadrant.*

The green quadrant shows responses that could be classified into one of our four main gender identity categories by both methods. The participants who were consistently classified have already been discussed.

In Study 1, 18 participants (3.9%) who were classified as one of the four main genders by either method were classified as a different one of the main four genders by the other method. Five participants selected Genderqueer for the multiple-choice measure but were classified as women by their free text responses, which were “butch leather dyke,” “butch female,” (indicated by two participants) “Butch,” and “I’m male, that’s my sex; I’m a girl, that’s my gender.” Two participants selected Genderqueer for the multiple-choice measure but were classified as agender by their free-text responses, which were “Agender” and “agender genderqueer femme faggot.” Three participants selected Man for the multiple-choice measure but were classified as non-binary by their free text responses, which were “Nonbinary trans man,” “Male, with an internal rarely shown but very real, if private identification as female, or of some complex mixed gender without a name but definitely non binary,” and “Genderqueer Trans Man.” Finally, eight participants selected Woman on the multiple-choice measure and were assigned to the non-binary category based on their free-text responses. Four of these responses were from trans women: “I’m a lovely nonbinary witch & a proud, angry trans woman,” “Non Binary Trans Femme,” “Genderfluid trans woman,” “Genderqueer trans woman.” The remaining four were “Genderqueer femme woman,” “Primarily female, but also genderfluid,” “Androgynous female-bodied person,” and “femme genderqueer.”

In Study 2, only five of the participants who were classified as one of the four main genders by either method were classified as a different one of the main four genders. Their free-text responses were as follows: “male,” “transgender female (femme),” “Agender/possibly male,” “agender,” and “primarily female.”

#### *Yellow Quadrant.*

The yellow quadrant shows participants who were classified as one of the four main genders by the multiple choice measure but not by the free text question.

In Study 1, 0.8% (N = 4) of participants were in this quadrant, and their free-text responses were “Bisexual,” “\*pterodactyl noises\*,” “I don't have gender feels, but as society treats me as a woman, I have a political gender identity as a woman,” and “Woman/nonconforming/has weird gendery thoughts sometimes.”

In Study 2, there were no participants in this quadrant (N = 0).

#### *Blue Quadrant.*

The blue quadrant shows participants who were classified as one of the four main gender identities by the free text measure but not by the multiple choice measure.

In Study 1, 3.2% (N = 15) of participants were in this quadrant. Five participants were inconsistently classified based on their responses to the multiple-choice measures, but classified as non-binary based on their free-text responses, which were as follows: “Sometimes I'm a lady and sometimes I have no gender,” “Genderqueer trans male,” “Genderfluid, typically leaning male,” “Transqueer nonbinary transboy,” and “Non-binary (agender).” One person who selected Choose Not to Answer for the multiple choice measure and one person who was inconsistently classified both wrote in “Male” for the free text response. The remaining eight participants had the following free-text responses: “Cis by default male,” “Woman, biologically and

psychologically but not behaviorally,” “Transfeminine / trans woman,” “Agender Trans woman,” “Non-binary,” “nonbinary/undefinable,” “None,” and “Genderfuck.”

In Study 2, 2.4% (N = 13) of participants were in this quadrant. Five of these participants wrote “male,” two wrote “female,” and one wrote “woman” as their free-text responses, and both the person who selected Don’t Know and the person who selected Choose not to Answer wrote in “Female.” The remaining five participants’ free-text responses were “male-brained in a female body, but not gender dysphoric,” “trans man,” “trans masculine,” “agender/male,” and “non-binary woman.”

#### *Red Quadrant.*

The red quadrant shows participants who could not be classified as one of the four main gender identities by either the multiple-choice measures or the free-text question

In Study 1, 4.3% (N = 20) of participants were in this quadrant. Eight of these participants were consistently classified by the two methods, but into categories that would probably be excluded from most samples, so I examined the free-text responses more closely to determine whether these responses could have been grouped into a category that is more likely to be used. Three participants selected “Another identity not listed” for the multiple-choice measures but could not be classified as one of our gender categories based on that response, and were similarly unclassified by their responses to the free-text measure (they were assigned to the Not Enough Information category). The responses for these participants were “7,” “Binary,” and “????????????????.” An additional five participants were categorized as Don’t Know by both the multiple-choice and free-text measures; these free-text responses were as follows: “I don’t know!,” “Dont know,” “Don’t know/female,” “Unknown,” and “Unknown, but some flavor of trans.” With the possible exception of the person who wrote “Don’t know/female,” who could

potentially have been grouped with women, it seems that these responses were too uninformative to be useable for most research.

In Study 1, the remaining 13 responses (2.8%) in the red quadrant were not consistent between the two methods. Two of these expressed uncertainty or confusion about their gender identities: “jfc you tell me i dont know man i dont have one,” and “Assigned female, masculine presenting most of the time, confused about my gender identity.” Three more indicated that gender did not apply to them (“Human, I don't care about gender,” “N/A,” “Gender irrelevant”). The remaining participants had the following responses: “I do not have one, because I do not identify with patriarchal gender roles,” “Without defining gender, I can't answer that,” “Weariness, or something like it,” “Autogynephillic,” “gender non conforming lezboi,” and “trans-fabulous bottom with a touch of raccoon.”

In Study 2, 0.9% (N=5) of participants were in the red quadrant. The one participant consistently classified as “Don’t Know” by both the multiple-choice and free-text question wrote “Questioning” as their gender in the free-text question, and one participant who could not be classified by either the free-text or multiple-choice questions wrote “I don't have one, I don't feel like gender should matter (though yes it does). I am somewhere in the spectrum of genders that is possible.” The remaining three participants indicated that their genders were “bisexual,” “I hate gender,” and “non-conforming.”

## **Comparison of Multiple-Choice Gender Identity and Gender Self-Categorization Measures**

### **Gender Self-Categorization (Joel et al. 2014).**

Table 18 shows the correspondence between responses to the multiple-choice gender identity question and the gender self-categorization question from Joel et al. (2014) for Studies 1 and 2. Based on theory, we expected that people who selected “Man” or “Woman” on the

multiple-choice gender identity question would also select “Man” or “Woman” respectively on the gender self-categorization question. We also anticipated that people who selected “Non-Binary” on the multiple-choice gender identity question would select “Both” or “Neither” on the gender self-categorization question, and that people who selected “Agender” on the multiple-choice gender identity question would select “Neither” on the gender self-categorization question.

In Study 1, 78.5% (N=212) of the 270 total responses to the gender self-categorization question were expected based on the responses to the multiple-choice gender identity question (shown in green in Table 18). Only 9.2% (N=25) were unexpected (yellow in Table 18).

In Study 2, 94.9% (N=550) of the 579 total responses to the gender self-categorization question were expected based on the responses to the multiple-choice gender identity question (shown in green in Table 18). Only 2% (N=12) were unexpected (yellow in Table 18).

These results suggest that the addition of Both and Neither options to the binary options captures additional information beyond the non-binary and agender options (because both non-binary and agender participants selected Neither most frequently). This is a limitation of the multiple-choice gender identity measure which should be addressed in future research. The difference between non-binary and agender identities is not straightforward, and there is substantial heterogeneity within each category. Future research should investigate the commonalities and difference between non-binary and agender identities, as well as any additional gender categories which may emerge in the near future.

Table 18

*Gender identity classification by multiple-choice and gender self-categorization (Joel et al., 2014) in Study 1 gender minority sample and Study 2*

Multiple-Choice Gender Identity	Gender Self-Categorization (Joel et al. 2014)							
	Man		Woman		Both		Neither	
	S1	S2	S1	S2	S1	S2	S1	S2
Man	40	235	-	-	2	1	2	3
Woman	1	-	43	271	3	2	3	-
Non-Binary	2	3	12	3	32	12	85	25
Agender	-	-	-	-	-	1	12	7
Not Enough Info	-	-	1	-	1	-	5	-
Don't Know	1	-	1	1	2	-	6	1
Choose Not to Answer	1	1	3	1	1	1	1	1
Inconsistent	-	6	3	3	1	1	6	-

Note. S1 = Study 1. S2 = Study 2.

Table 19

*Gender identity classification by multiple-choice measure and gender self-categorization (adapted from Rockquemore & Arend, 2002) in Study 2 (N=584).*

Multiple-Choice	Gender Self-Categorization (adapted from Rockquemore & Arend, 2002)							
	Man <sup>a</sup>	Woman <sup>b</sup>	Binary/ NB <sup>c</sup>	NB, Man <sup>d</sup>	NB, Woman <sup>e</sup>	NB <sup>f</sup>	Rejects Gender <sup>g</sup>	Something Else <sup>h</sup>
Man	222	1	2	2	-	1	8	2
Woman	-	257	3	1	7	-	3	5
Non-Binary	1	-	6	1	17	9	1	8
Agender	-	-	-	-	2	2	-	4
Don't Know	-	-	-	-	1	-	2	-
Choose Not to Answer	1	1	-	1	-	-	-	-
Inconsistent	5	5	-	-	1	-	-	2

<sup>a</sup>I consider myself exclusively a man.

<sup>b</sup>I consider myself exclusively a woman

<sup>c</sup>I sometimes consider myself a man or a woman, and sometimes non-binary depending on the circumstances.

<sup>d</sup>I consider myself non-binary, but I experience the world as a man.

<sup>e</sup>I consider myself non-binary, but I experience the world as a woman.

<sup>f</sup>I consider myself exclusively non-binary (neither a man nor a woman).

<sup>g</sup>Gender is meaningless. I do not believe in gender identities.

<sup>h</sup>Something else (please describe).



### **Gender Self-Categorization (adapted from Rockquemore & Arend, 2002).**

Table 19 shows the correspondence between responses to the multiple-choice gender identity question and the gender self-categorization question I developed based on Rockquemore and Arend (2002). Based on theory, we expected that most people who selected “Man” or “Woman” on the multiple-choice gender identity question would also select “Man” or “Woman” respectively on the gender self-categorization question. However, we expected that some others would choose a category reflecting a dislocation between identity (as non-binary) and practical experience (as a man or woman), or a category reflecting vacillation between multiple categories (as binary and non-binary). A majority of participants (87.6%) were classified the same way between both questions, and only two participants gave fully inconsistent responses (red cells in Table 19). Most participants were classified as exclusively men or exclusively women on both measures (N = 479, or 82%), and 5.6% (N=33) of participants indicated that they identified as non-binary on both measures (green cells in Table 19).

Another 2.3% of people (N=14) who identified as men or women on the multiple-choice gender measure indicated that they identified as non-binary but experience the world as a man or woman, or that they vacillate between identifying as binary and non-binary (blue cells in Table 19). One possible explanation for this result (in line with Rockquemore and Arend, 2002) is that lack of social recognition of one’s identity category may reduce the likelihood of claiming or disclosing that identity. That is, when non-binary gender identities are not widely socially recognized as *possible*, and are not given social reality, claiming them may feel extremely difficult (even when given the option in a survey question). This problem is distinct from the one faced by binary transgender people; while many face identity invalidation and denial in terms of their membership in their identity, no one doubts the existence of the categories “women” and

“men.” Future research is needed to investigate the distinct experiences of non-binary people, and what factors in their social environments influence their labelling of their identities, particularly the relationship between social recognition of a group’s existence and strength of identity with that group. These interesting possibilities aside, while these results suggest a moderate level of construct validity for the multiple-choice gender identity measure, there is clearly more diversity present within categories (especially within the non-binary category). In addition, about 10% of the sample (N=58) selected an unexpected combination of responses. Twenty-one participants chose to write in a different response; these answers are shown in Table 20, alongside the participants’ multiple-choice gender identity and transgender identity.

The responses suggest that participants are extremely aware of social norms around gender categories and that many of them are actively determining how to make/find space that fits them (e.g. *“Lean heavily towards being a woman, not sure what else there is yet.”* or *“I consider myself to have aspects of men and women but I don't really fit either. I am my own thing.”*). Several participants (especially agender people) emphasize that gender is relatively unimportant/meaningless to them but not to others. For example, one person describes gender as *“a lens we force on others, not how many of us actually are.”* Still other participants suggest that strength of identification with gender categories is an important dimension of their experience:

“I experience aspects of male and female identification, sometimes at the same time, but neither of them is a strong or complete identification (my identification with them is always partial and weak, so neither label fits well even though both have elements of the truth).”

Taken together, these responses suggest several avenues for future research about the gender categories which participants perceive as available to them, the categories they feel others assign to them, and the strength of their identification with any of these available categories (see van Anders, 2015 for a recent thought-provoking approach to these questions).

Table 20

*Free-Text Comments on “Something Else” Choice for Gender Self-Categorization Measure (adapted from Rockquemore & Arend, 2002) in Study 2.*

Gender Identity	Trans Identity	Comment after Selecting “Something Else”
M	T	I don't consider gender particularly meaningful, but am aware of it as a social norm. I am mostly perceived as a woman but would like to be able to be perceived as a man.
M	T	I consider myself male and identify as male, and am perceived by society to be a man based on my behavior and presentation (e.g. clothes, mannerisms, hair,), but also feel very non-binary because 22 years of acting like and being perceived like a girl/woman has influenced how I experience the world.
W	T	Lean heavily towards being a woman, not sure what else there is yet.
W	T	I consider myself a trans women. It's different than just being a "women" but I also believe that gender shouldn't be labeled. As a society we should just let people be people.
W	T	I consider myself a trans-woman, with the "trans" part very much at the fore at the moment.
W	NT	I identify as a woman and perceive myself as a woman, but am more androgynous, without being non-binary.
W	NT	I consider myself a woman because it would not be possible for me to be a man in the full biological sense, but if I could I would be reborn male but able to adjust myself to also look female
NB	T	mostly non-binary, but sometimes a woman.
NB	T	I experience the world as non-binary, but the world perceives and reacts to me as a woman
NB	T	I experience aspects of male and female identification, sometimes at the same time, but neither of them is a strong or complete identification (my identification with them is always partial and weak, so neither label fits well even though both have elements of the truth).
NB	T	I'm a prettyboy but people treat me as a woman
NB	T	I consider myself non-binary, and experience the world as a person who is a combination of masculine and feminine energies but am often perceived and treated as a man.
NB	NT	I consider myself to have aspects of men and women but I don't really fit either. I am my own thing.
NB	NT	A man with feminine characteristics
NB	DK/NA	I consider my soul agender, but I experience the world as a woman.
A	T	None
A	T	I do not understand the concept of gender but understand it is not that way for everyone
A	T	Gender has meaning to others, so it can't be meaningless, but I believe that gender is a lens we force on others, not how many of us actually are. I don't identify with either gender, but I respect the pride others can take in it.
A	DK/NA	I don't identify as any particular gender, but that doesn't mean I think gender is meaningless to others
I	DK/NA	i hate it but more or less "sometimes man or woman sometimes nb"
I	-	i consider myself pretty much evenly a man and nothing at the same time most of the time, sometimes i just feel like nothing, sometimes i just feel like a man, i never ever feel like a woman

*Note.* M: Man, W: Woman, NB: Non-Binary, A: Agender, I: Inconsistent. T: Transgender, NT: Not Transgender, DK/NA: Don't Know/No Answer

## Comparison of Transgender Status Measures

We assessed transgender status with three different measures in Study 1 (transgender identity, transition history, and gender identity + sex assigned at birth) and two different measures in Study 2 (transgender identity and gender identity + sex assigned at birth). For each question, participants were categorized as Not Transgender, Transgender, or Unclassified as described in the Methods. The number of participants in each category for each measure is shown in Figure 12 for Study 1 and Figure 13 for Study 2.

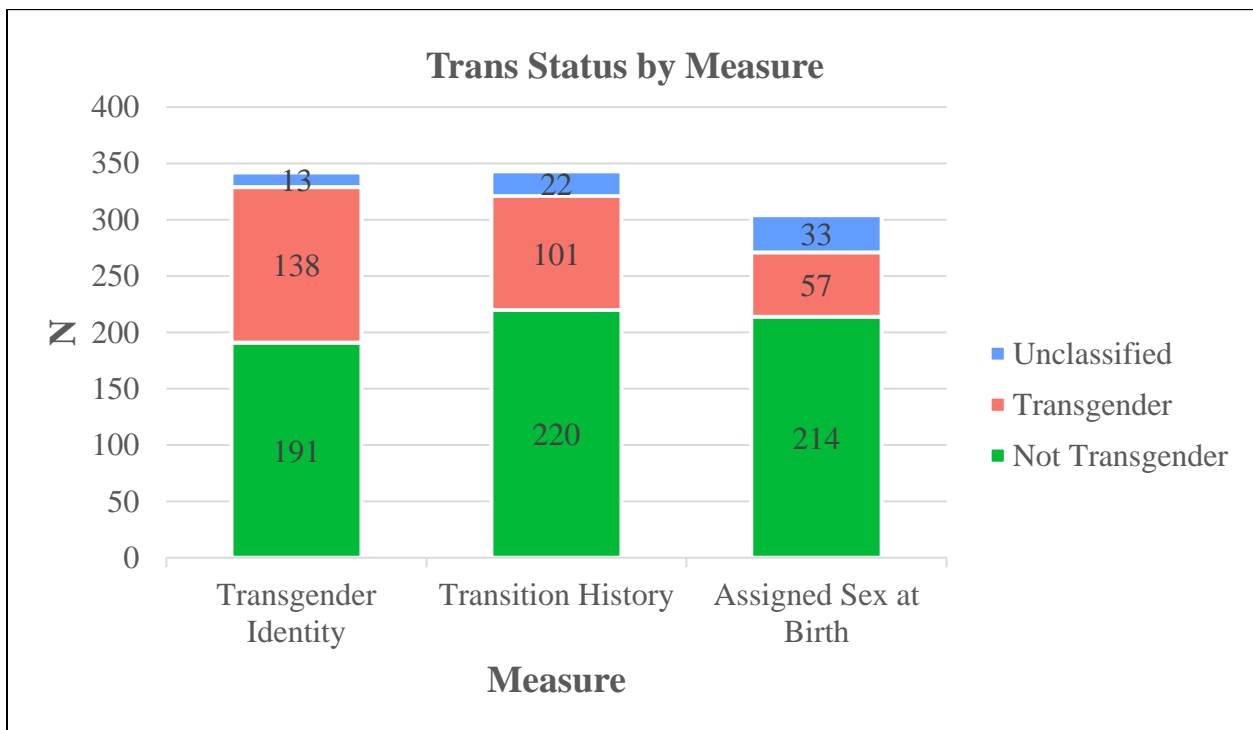


Figure 12. Transgender status of participants by measure in Study 1.

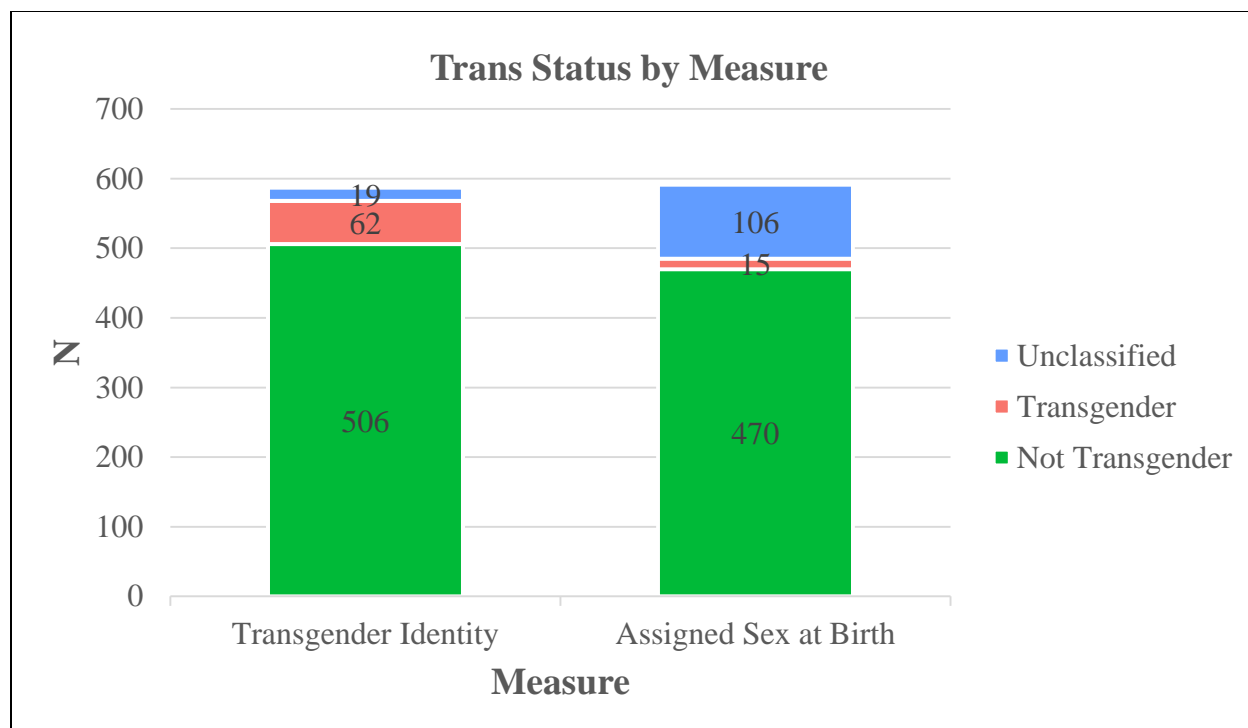


Figure 13. Transgender status of participants by measure in Study 2.

The number of Unclassified participants is consistently lowest for the Transgender Identity question (and Transition History, in Study 1), because that category includes only participants who selected “Don’t Know” or “Choose not to answer” when asked whether they identify as transgender/whether they have a history of transition. In contrast, for the Assigned Sex at Birth question, all participants who selected a gender identity other than “Man” or “Woman” or an assigned sex at birth other than “Male” or “Female” (e.g. all non-binary and agender people) appear in the Unclassified category. Finally, the number of Unclassified people for the Free Text question is so large because the question did not explicitly ask participants to report transgender status, and all participants who did not explicitly describe themselves as cisgender or transgender (e.g. people who wrote “man” or “woman”) were categorized as Unclassified.

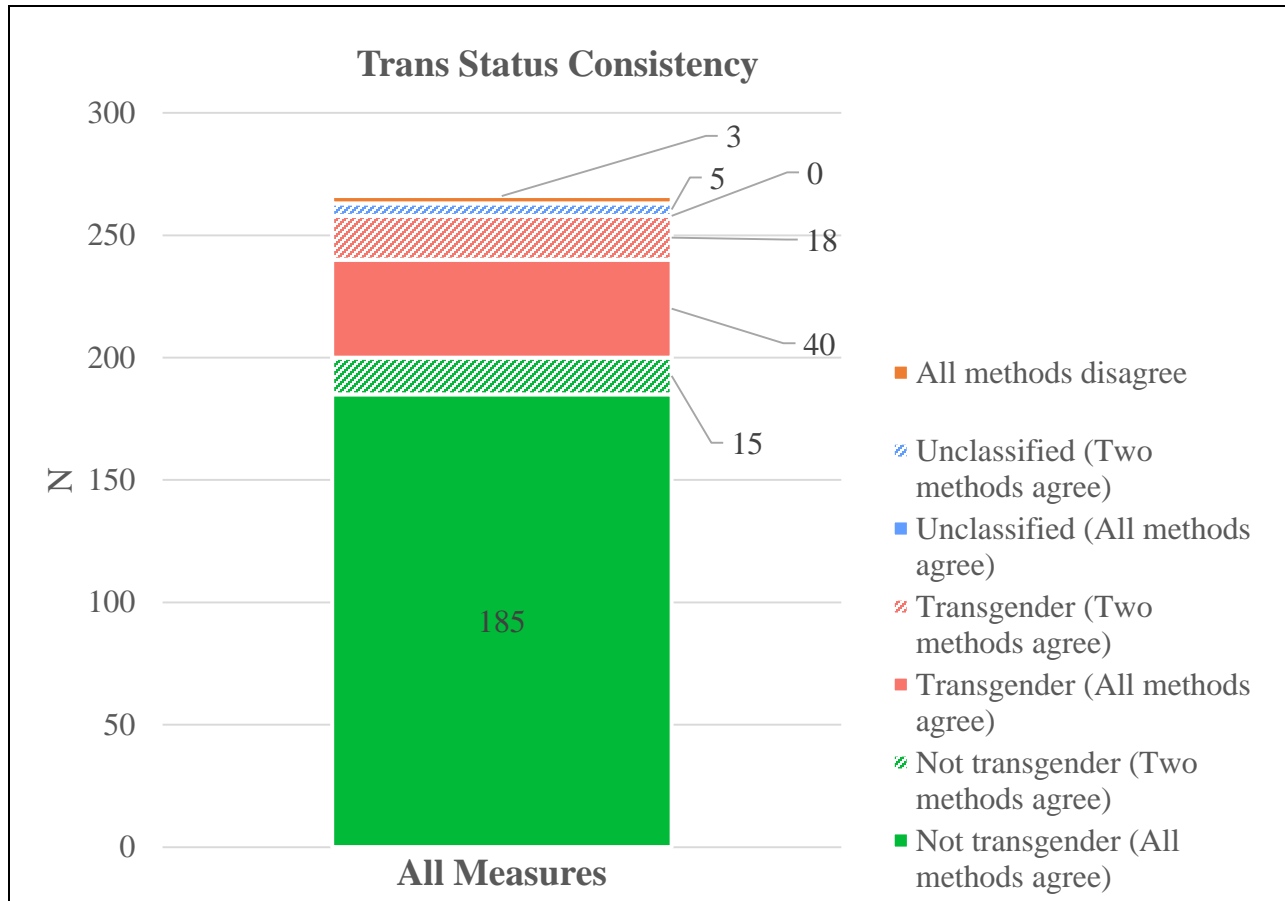


Figure 14. Consistency of transgender status classification across transgender identity, transition history, and assigned sex at birth measures (Study 1).

Figure 14 shows the consistency of transgender status classification across the three measures in Study 1. Not all participants answered all three measures, so these comparisons are based on the 270 participants who did complete all three measures. Eighty-three percent (N=225) were consistently classified as Not Transgender, Transgender, or Unclassified by all three measures (solid bars in Figure 14), while 15.5% (N = 42) were classified the same way by two out of three measures (striped bars in Figure 14). Three participants were classified fully inconsistently (i.e. differently by each of the three measures).

Table 21

*Transgender status classification by transgender identity and transition history measures (Study 1; N=342)*

<b>Transition History</b>	<b>Transgender Identity</b>		
	Not Transgender	Transgender	Unclassified
Not Transgender	<b>237</b>	25	6
Transgender	3	<b>47</b>	3
Unclassified	-	17	<b>4</b>

Table 22

*Transgender status classification by transgender identity and gender identity + sex assigned at birth measures (Study 1; N=271)*

<b>Gender Identity + Sex Assigned at Birth</b>	<b>Transgender Identity</b>		
	Not Transgender	Transgender	Unclassified
Not Transgender	<b>187</b>	10	11
Transgender	2	<b>43</b>	-
Unclassified	-	18	-

Table 23

*Transgender status classification by transition history and gender identity + sex assigned at birth measures (Study 1; N=270)*

<b>Gender Identity + Sex Assigned at Birth</b>	<b>Transition History</b>		
	Not Transgender	Transgender	Unclassified
Not Transgender	<b>197</b>	7	4
Transgender	1	<b>41</b>	2
Unclassified	1	12	<b>5</b>

Table 24

*Transgender status classification by transgender identity and gender identity + sex assigned at birth measures (Study 2; N=523)*

<b>Gender Identity + Sex Assigned at Birth</b>	<b>Transgender Identity</b>		
	Not Transgender	Transgender	Unclassified
Not Transgender	<b>458</b>	3	14
Transgender	2	<b>13</b>	-
Unclassified	-	30	<b>3</b>

As shown in Tables 21-23, the rate of agreement about transgender status between the transgender identity and transition history measures was 84.2%, the rate of agreement between transgender identity and gender identity + sex assigned at birth was 84.8%, and the rate of agreement between transition history and gender identity + sex assigned at birth was 90%. None of these categorization methods perfectly overlap with any of the others.

Table 24 shows the consistency of transgender status classification between the two measures for Study 2. Ninety percent (N= 474) were consistently classified as Not Transgender, Transgender, or Unclassified by both measures. These results suggest that there is substantial overlap in the results produced by these two measures.

## **Discussion**

### **Gender Identity Measures**

#### **Binary vs. Multiple-Choice.**

As was theoretically expected based on the literature presented in Chapter 1, the multiple-choice gender identity measure was able to identify substantially more diversity in the same sample than the binary measures. Between 7.3% (in Study 2, a majority cisgender sample) and 34.3% (in Study 1, a majority gender minority sample) of participants categorized as men or women by the binary question were classified as a different gender by the multiple-choice question.

#### **Binary + Other vs Multiple-Choice**

Correspondence between the binary + other measure and the multiple-choice question was fairly high (85% the same in Study 1, 95% the same in Study 2) and on par with the free-text question. If we consider only this evidence, binary + other question performs about as well as most others, but consistency of responding is not the only consideration.



The fact that we recategorized several participants who had selected “Other” back into the Man or Woman categories based on their free-text responses (23 people in Study 1, four in Study 2, all of them transgender people) was somewhat concerning given our theoretical expectation that the Binary + Other coding cues transgender participants to believe that researchers do not regard their gender identities as legitimate (e.g. Pope & Warner, 2017). This possibility is explored in more depth in Chapter 4, where we conclude that the binary + other question is disliked by gender minorities and cisgender people alike, and may act as a cue to identity threat for gender minorities. Considering that it requires nearly as much hand-coding of responses as the free-text question, and is dramatically less well-liked, we do not recommend it.

#### **Free Text vs. Multiple-Choice.**

The number of responses that were classified differently by the multiple-choice and free-text responses was much higher in Study 1 than Study 2. This is presumably because the proportion of gender minority participants was much higher in Study 1, and we might expect that the responses from this group are genuinely more complex to categorize. In addition, the absence of an explicit Agender category in the multiple-choice question in Study 1, and the labeling of one option as Genderqueer rather than Non-Binary, may help explain the relatively higher proportion of inconsistency in Study 1.

Considering the results from both studies, it seems that responses to the free-text question could be meaningfully categorized somewhat more often than responses to the multiple-choice question (as demonstrated by the consistently larger numbers of participants in the blue quadrants than the yellow quadrants). This was true despite the fact that the multiple-choice question includes an option to write in a free-text response, which was coded and categorized by hand using the same coding scheme. It may be that participants believed that researchers were

interested in a greater level of detail when the free-text box was the only gender measure. In contrast, they may have assumed (often accurately) that write-in responses to the multiple-choice measure were more likely to be discarded as unusable, and thus did not put in as much effort to clearly express their identities in that context.

Many of the differences between free-text and multiple-choice categorizations seem to originate with participants whose experiences of gender cannot be neatly captured by the available options in our multiple-choice question. Ironically, the process of coding the free-text responses is more able to impose a rigid categorization structure on the experiences of participants who are unable or unwilling to impose those categorization structures on themselves. For example, consider the participant in Study 2 who selected “Genderqueer” for two multiple-choice questions and “Woman” for the other two. This participant wrote in “Agender trans woman” for the free-text response, which we coded as “Agender” because our coding guidelines opted to prioritize agender identity when multiple identities were written in at once. However, it is at least equally defensible to code this response as “Woman.” Given the explicit choice between the two, this participant quite reasonably chose both equally often.

Other participants were classified inconsistently by the free text and multiple-choice measures because the researchers’ interpretation of the meaning of a given gender label differs from the participants’ own interpretation of that label. For instance, in Study 1, four participants selected “Genderqueer” as their gender identity on the multiple-choice measure and wrote “Butch” for the free-text measure. According to our coding scheme, “Butch” is a subset of the gender category “Woman,” so it was grouped accordingly. However, given that the participants themselves indicated that they both use the label “Butch” and feel the category “Genderqueer”

best reflects their understanding of their gender identities, the coding scheme is clearly invalid for this case.

Both the free-text coding scheme and multiple-choice measure are also limited in their ability to accurately capture the experiences of people whose gender spans more than one of our designated categories (e.g. “Non-binary woman,” “Agender Trans woman,” “Non-binary (agender)” or fluctuates (e.g. “Genderfluid trans woman,” “Genderfluid, typically leaning male”). In the free-text coding scheme, we gave priority to the underrepresented categories “non-binary” and “agender” when multiple gender labels appeared in a response, but participants frequently selected a different category for themselves on the multiple-choice question. It seems likely that if the multiple-choice measure permitted multiple selections, many of these participants would have chosen more than one (indeed, some recent research found that up to 90% of gender minority participants selected multiple options when given the opportunity; Staples, Bird, Masters, and George, 2018). While researchers would probably better capture their experiences by allowing them to do so, the difficulty of analyzing multiple-select questions make this approach too complex to recommend for general use.

Given all these concerns, we find that any method which requires coding all responses or substantial simplification of data from a multi-select form (e.g. choosing which ones to prioritize, or creating dozens of different categories) presents too much of a barrier to entry for general use. The development and maintenance of a valid coding scheme for free-text responses would require multiple rounds of cognitive interviewing with participants, with frequent changes likely being required, as gender identity terminology, usage, and visibility have been changing rapidly in recent years (e.g. the shift from genderqueer to non-binary as an accepted umbrella term). Likewise, most common statistical techniques require reducing the data from multi-select

questions to mutually exclusive categories before analysis, adding work for researchers. Even with this level of attention to detail and investment of resources from researchers, the addition of these extra steps reduces the likelihood that the ultimate categorization of a participant is one they would agree with, thus reducing the validity of the data these measures produce.

Instead, we conclude that the difficulties of categorizing a wide range of participants into a few gender identity groups are best left to the participants themselves, i.e. by presenting them with single-select multiple-choice measures. Researchers should invest their time and energy into informing their participants about the purposes for which their information will be used and the reasons why the categories in the multiple-choice question were selected, emphasizing the fact that statistical analysis limits them to offering only a few options. If a free-text response option is included (as we recommend), it is a built-in mechanism for identifying options which should be modified and additional categories which should be added (for example, we added Agender as an option and changed the Genderqueer option to Non-Binary in Study 2 on the basis of free-text responses participants wrote in for Study 1). Researchers might also explicitly inform participants that if they do not choose one of the main provided options and write in another, their data will be recategorized by the researcher or excluded from analysis. While this might artificially increase homogeneity of responding, this outcome is likely preferable to the loss of data from participants whose responses would otherwise not be useable. This approach leads to two highly positive outcomes; it reduces work for researchers while increasing the level of control given to participants to self-determine how they are categorized (and thus the validity of their data).

## **Transgender Status Measures**

As described for gender identity measures, inconsistency of classification between transgender status measures is partially the result of researcher choices about how to classify particular response patterns, which may or may not align with the choices participants would make for themselves. In particular, the decision to classify all non-binary and agender people as Unclassified for the Assigned Sex at Birth measure likely caused much of the inconsistency (the substantive theoretical and pragmatic question about whether non-binary and agender people are transgender, or transgender in the same way as binary transgender people, remains outside the scope of this paper).

Another cause of inconsistency, however, is much more pronounced for transgender status measures than for gender identity measures. While the multiple-choice and free-text gender identity measures are both still assessing gender identity, these transgender status measures are actually assessing theoretically distinct constructs (see Chapter 2 for detailed explanations). A person whose gender identity differs from their sex assigned at birth, which is a concrete (if invasive) question about their birth certificate, does not necessarily identify as transgender, which is among other things a question about group membership and perception of oneself as belonging to the social category of “transgender people.” Likewise, a person who has a history of transition (social or otherwise) need not identify as transgender, and vice versa. This is why Balarajan et al. (2011) recommend that all three of these measures be included in surveys aiming to maximize their identification of transgender participants, a recommendation we echo for researchers in that situation. However, most research does not require this level of exhaustive identification of transgender participants, which is unnecessarily lengthy, invasive (see Chapter 4), and complex to consolidate for analysis.

Overall, the Assigned Sex at Birth question seems to work as intended to identify binary transgender people, and may be useful if information about physical bodies is needed (e.g. for health research). It is less well-suited for psychological research, as it does not provide any additional information about the transgender identities of non-binary and agender people. An additional drawback of this method is the complexity of analyzing its results, which requires comparison of responses to two measures to determine transgender status. In contrast, responses to the Transgender Identity and Transition History questions are relatively simpler to analyze and can be interpreted in the same way regardless of a participant's gender identity, minimizing researcher labor and maximizing the number of usable responses.

### **General Recommendations**

In both samples, the multiple-choice measure captured much more gender diversity than did the binary question, about the same amount as the hand-coded binary + other question, and somewhat less than hand-coded free text responses. We do not recommend the binary + other question because it requires a substantial effort to hand code responses, is not well-liked, and may produce identity threat for gender minorities by literally “othering” them (see Chapter 4). In contrast, the free text measure is well-liked and non-threatening, and will certainly produce meaningful and nuanced categorizations, but because of the dramatically increased time required to code the data it produces, and the loss of self-determination for the participants, we do not recommend this method for general use. Instead, we recommend the multiple-choice measure, because it requires much less time and effort for researchers, and gives more control to participants.

For general use, we find that the transgender identity question and transition history question are equally easy to administer and analyze because they can be used as-is without the

recoding which is necessary for the Sex Assigned at Birth question. The transgender identity question has our final recommendation because it has greater face validity for concepts related to identity (which is of greater relevance for social psychology) and does not invoke the ambiguities and issues related to medicalization and gatekeeping of transgender people's identities that are raised by asking about transition history (see Chapter 4).

## Chapter 4: Acceptability and Identity Safety of Gender Measures

In the previous chapter, we considered how to analyze the results of several different transgender-inclusive gender measures, with the aim of identifying the questions that best balanced accuracy and simplicity of use. This chapter explores whether these same gender measures may act as cues to identity safety for transgender and non-binary participants, and how these cues may differ across various contexts in which researchers might use gender measures. Our aim is to identify the gender measures that are most acceptable to participants of varying gender identities and transgender statuses, and that provide the greatest sense of identity safety across contexts.

### Identity Threat and Identity Safety

Several researchers who have conducted qualitative research about inclusive gender questions have pointed out that many transgender people are uncomfortable answering certain questions or do not feel that they can adequately express their identities using the options provided (e.g. Conron et al., 2008; Balarajan et al., 2011; Fenway Health, 2013). Despite the quantity of qualitative research suggesting that transgender people may react negatively to certain gender question formulations, no one has yet connected these findings to social psychological research on *identity threat* and *identity denial*.

*Identity threat* is the recognition that one may be devalued in a given setting because of one's social identity. Social psychological research on this subject suggests that members of stigmatized groups are particularly sensitive to environmental cues that may help them determine whether identity threat is likely in a given setting. As Murphy, Steele, and Gross (2007) describe, “when a setting contains threatening situational cues, it raises the specter of identity threat—prompting heightened cognitive and physiological vigilance, decreased feelings of belonging,



and decreased desire to participate in the setting” (p. 884). While they are not an “environment” in the conventional sense, it seems clear that the type and wording of gender questions might serve as cues to identity threat or identity safety for gender minority participants in various settings (e.g. job applications, psychological experiments, health care clinics, social media sites).

In addition, some kinds of gender questions may even act as *identity denial* for gender minority participants. This research has largely investigated participants’ reactions to verbal questioning of their racial or national identities. For example, Cheryan and Monin (2005) found that Asian-American participants (but not white Americans) worked harder to demonstrate American cultural knowledge (reasserting American identity) when asked whether they spoke English before the experiment. While the question may seem benign in isolation, repeated questioning of this type (e.g. “where are you really from?”) can reinforce the idea that marginalized groups do not fully belong. In many contexts, the targets of this threat may react with *identity assertion*, redoubling their efforts to demonstrate membership in the group, as the participants in Cheryan and Monin (2005) did.

However, answering a demographic question alone does not provide much opportunity for identity assertion, especially when the question itself is the source of identity denial. In this situation, threatened participants may be more likely to withdraw effort from the task at hand. Townsend and colleagues (2009) proposed that multiracial participants who are forced to choose only one racial category on a racial demographic question may experience identity denial, reasoning that participants who hold more than one racial identity will inevitably feel that their identities are being denied if they are only allowed to communicate one of those identities to the researchers. Indeed, multiracial participants forced to select only one racial identity option subsequently showed lower motivation and self-esteem than multiracial participants who were

allowed to select more than one option. Townsend et al. (2009) argue that these effects are evidence of self-induced identity denial experienced when completing forced-choice racial identity questions.

Both identity threat and identity denial may be at work when gender minority participants encounter certain types of gender questions. For instance, when genderqueer or other non-binary identified people encounter a binary gender question, they may experience identity denial, as they can anticipate that the researchers will misperceive their identities. Likewise, if a transgender man sees himself as both transgender and a man, he may experience identity threat and identity denial if a gender question forces him to choose between these identities. By including a “transgender” option as a gender identity (e.g. in a set of responses like man/woman/transgender) the researchers have signaled a belief that transgender people cannot be men or women, and that they are likely to misperceive and devalue this participant’s identity as a transgender man.

### **Consequences of Identity Threat/Identity Denial**

To the extent that certain gender questions induce identity threat and/or identity denial, previous research suggests that gender minorities will feel lower self-esteem and motivation, decreased investment in the task at hand, reduced willingness to disclose sensitive information, reduced goodwill toward the person asking the question in general, and have lower rates of participation.

These effects are particularly concerning for psychological research because gender minorities are often already poorly represented in psychological research and mistrustful of researchers. As the APA Task Force on Gender Identity and Gender Variance (2009) noted,

“some transgender people, particularly some community activists, have been disillusioned by traditional research. ...The clinical language, the inclusion of some types

of gender variance in the DSM, the apparent focus on prevention, and the perceived gatekeeping role of some research are alienating and stigmatizing” (p. 26).

Though not all psychological research is clinical in focus, this distinction may not be apparent to non-psychologists, and their concerns are likely to carry over to other areas of psychology as well.

As the psychological community has not historically had the best interests of gender minorities in mind, many gender minorities are understandably wary of the intentions of psychological researchers. However, the same participants may help researchers to recruit additional participants if they perceive the studies as inclusive and considerate. For instance, Meier and Labuski (2013) have found that

“because the trans community has a strong Internet presence, they are well connected and can refer many other trans people to studies that they deem ‘sensitive’ and ‘worthwhile,’ whereas they may also warn others not to participate in studies not considered ‘safe’” (p. 301).

Both the preexisting barriers to the participation of gender minorities in psychological research, and the potential these same people have to increase participation of the other members of their communities make maintaining the goodwill and trust of gender minority participants a crucial goal for psychologists. Identifying which questions serve as cues to identity safety and acceptance and which serve as cues to identity threat and denial may be a major help in accomplishing that goal.

The existing literature on trans-inclusive gender questions suggests that merely allowing participants to choose multiple gender identities or to identify themselves as transgender may not be sufficient to reduce identity threat or identity denial. I propose that participants with a history of transition who do not identify as transgender (but who nevertheless may be viewed as transgender by the researchers) may experience similar feelings of identity denial when they

encounter the two-question method and are asked to disclose their sex assigned at birth. Most transgender people are aware of the significance of disclosing different information for sex assigned at birth and gender identity, and know that doing so will cause researchers to identify them as transgender. Thus, questions that ask participants to report their sex assigned at birth may cause both identity threat and identity denial, as they suggest that researchers either do not understand or do not value participants' own gender identities (or both).

This perspective may help to make sense of the actions of the participants in Glen and Hurrell (2012) who report a history of transition but do not indicate a difference between sex assigned at birth and current gender identity. To the extent that they do not feel their sex assigned at birth accurately describes their identity, these participants might be acting to avoid self-inflicted identity denial by reporting their sex assigned at birth as congruent with their gender identity. For participants who do not regard their sex assigned at birth as relevant to their gender identity in any way, it is even possible that merely asking about sex assigned at birth will induce the experience of identity denial. Consistent with these ideas, Conron et al. (2014) seem to find that some participants are more comfortable answering a question about transgender identity than about sex assigned at birth. Likewise, Glen and Hurrell (2012) find that some participants seem more comfortable indicating that they have a history of transition than indicating that their sex assigned at birth is different from their current gender identity.

The following set of analyses examines the possibility that different formulations of inclusive gender questions may pose different risks of identity threat. We hypothesize that the level of identity threat or identity denial experience will vary along with individual differences in participants' identities. For example, some binary transgender people may not feel identity denial when they see the binary gender question, as they are able to indicate their genders. Participants'

philosophies about the nature of those gender identities may also influence their experiences of threat. Some non-binary people who view their identity as occupying a “third space” beyond man and woman may not experience identity threat when they see “other” as a third option in addition to “man” and “woman,” because this question format conforms to their own philosophy of gender. Finally, we propose that these varying levels of perceived risk may partially account for between-participant differences in willingness to disclose information and participate in research.

### **Effects of Context**

Nearly every organization that has published recommendations about gender questions also notes that it is crucial for respondents to know what the information is going to be used for, both so that they can respond accurately and so that they have trust in the people asking the question. For instance, the Equality and Human Rights Commission (2012) suggests that

an introductory text should be provided before this question in order to encourage participation and to gain respondents’ trust to answer the questions. The text should explain *why the data are being collected; how the data will be protected...* and *how respect for the respondents’ answers will be achieved by the organization* [emphasis mine]; for example, through ensuring confidentiality and anonymity throughout (p. 12).

This suggests that the context in which gender questions are asked may have significant effects on participants’ willingness to answer the questions in the first place.

It also seems possible that the way participants think about these identities and label themselves may differ according to context. In particular, I propose that contexts vary in terms of the *gender salience* and *motivation to be perceived accurately* that they elicit. That is, gender is likely to be a highly salient identity category in some contexts (e.g. when searching for a romantic partner) and much less salient in others (e.g. during grocery shopping). Likewise, in some contexts people are highly motivated to ensure that others perceive their gender identities

accurately (e.g. when on a date), while in other circumstances it is less crucial to be perceived accurately (e.g. when at a movie theater alone). It may be that the context in which gender questions are asked will influence participants' comfort with the questions and perceptions of the research and researchers.

## Method

Like Chapter 3, this chapter describes results from different parts of the same two studies (each sampling undergraduates and self-identified gender minorities for four total samples). For a detailed description of the sample, the gender demographic measures, and the coding techniques used for each one, see Chapter 3. For a detailed description of each question and how it was chosen, see Chapter 2.

### Design Overviews

In Study 1 (Spring 2016), we used a fully-crossed 3 (context) x 4 (gender question) design to determine the effects of survey context and gender question type on participants' expectations of identity threat and identity denial. To manipulate context, we randomly assigned each participant to see one version of the following paragraph:

Please answer the following gender question(s) as you would if you were completing a demographic questionnaire **on a clinic intake form/for a psychology experiment/on an online dating website.**

These contexts were chosen because gender questions are likely to be asked in each of them, but the researchers and participants are likely concerned with different components of gender in each circumstance (e.g. asking about physical bodies via sex assigned at birth may be viewed as more relevant and acceptable in a health care context than in a psychology experiment). After reading a context description, participants were randomly assigned to see one of four different gender questions, designed to represent the most common types of inclusive gender questions currently

recommended by major public health and demographic research organizations. These were the Binary plus Other, Multiple-Choice, Multiple-Choice + Sex Assigned at Birth, and Multiple-Choice + Transgender Identity.

In Study 2 (Fall 2016), we did not include a context manipulation, which permitted us to experimentally compare more gender questions while reducing the number of cells in the design to increase power. We kept the same measures which were compared in Study 1, except that we removed the multiple-choice + transition history question because it was poorly ranked, and added the binary measure to allow direct comparison of the gender minority inclusive questions to the status quo. We also added the short free-text question, which was consistently viewed most positively by participants in Study 1. Thus, participants in Study 2 were randomly assigned to see one of six gender demographic measures (Binary, Binary + Other, Multiple-Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, or Free Text).

In both Study 1 and Study 2, after completing the randomly assigned gender measure, participants completed several dependent measures intended to assess their levels of identity threat, perceptions of the researchers conducting the survey, and their investment in future research. In Study 1, items were randomized within scales, but scales were presented in the same order each time. In Study 2, items were randomized within scales and scales were presented in a random order within blocks of the survey. Participants then completed all the same gender identity and transgender status measures that were randomly assigned at the beginning of the study, provided feedback about each one, and rated their level of preference for each one. In Study 1, participants also completed the Binary, Multiple-Choice + Transition History, and Long Free-Text measures at the end of the survey, although these were not part of the randomly assigned pool at the beginning of the survey. Thus, each participant completed one randomly

assigned gender measure at the beginning of the survey, and eight (in Study 1) or six (in Study 2) gender measures at the end.

## **Experimental Effects of Question Type (Between-Participants Manipulations)**

### **Identity Threat Measures.**

#### ***Study 1.***

After the experimentally assigned gender measure, we measured gender salience.<sup>16</sup>

Participants were then prompted to write free-text responses giving feedback on the measure (what they liked, what they didn't like), perceptions of the organizations that created the measure, and any additional comments. Participants then answered several questions designed to measure various aspects of identity threat, worded to match the context manipulation to which they were assigned. Participants indicated their answer to each question on a 7-point Likert scale (1 = Not at all, 7 = Very much).

*Being Understood and Respected* ( $\alpha = 0.90$ ,  $M = 5.0$ ,  $SD = 1.8$ ).

- *How likely do you think it is that [the clinic staff/ others on the dating website/ the experimenters] will **accurately perceive** your gender?*
- *How likely do you think it is that [the clinic staff/ others on the dating website/ the experimenters] will **correctly understand** your gender?*
- *How likely do you think it is that [the clinic staff/ others on the dating website/ the experimenters] will **respect** your gender?*
- *How likely do you think it is that [the clinic staff/ others on the dating website/ the experimenters] will **accept** your gender?*

*Comfort with Question* ( $\alpha = 0.93$ ,  $M = 5.7$ ,  $SD = 1.4$ )

- *How **comfortable** do you feel answering this gender question?*
- *How **comfortable** would you feel knowing that [the clinic staff/ others on the dating website/ the experimenters] had this information about you?*
- *How **safe** do you feel answering this question?*

---

<sup>16</sup> These questions were “How much are you currently thinking about your gender?” and “How much are you currently focusing on your gender?” Unfortunately, because these questions appeared immediately after the gender question manipulation, some participants seem to have believed that they were part of the gender question manipulation, and their comments reflected this confusion. Because of this confusion, these two questions were not included in any additional analyses.



- How **safe** would you feel knowing that [the clinic staff/ others on the dating website/ the experimenters] had this information about you?

Willingness to Participate ( $\alpha = 0.93$ ,  $M = 5.1$ ,  $SD = 1.6$ ).

- How **comfortable** would you feel [visiting this particular clinic/using this particular dating website/ participating in this particular psychology experiment]?
- How **safe** would you feel [visiting this particular clinic/using this particular dating website/ participating in this particular psychology experiment]?
- How **willing** would you be to [visit this particular clinic/use this particular dating website/ participate in this particular psychology experiment]?
- How **much** would you like to [visit this particular clinic/use this particular dating website/ participate in this particular psychology experiment]?

Willingness to Refer a Friend ( $\alpha = 0.95$ ,  $M = 4.7$ ,  $SD = 1.8$ ).

- How **willing** would you be to refer a friend to [this particular clinic/this particular dating website/this particular psychology experiment]?
- How **enthusiastically** would you refer a friend to [this particular clinic/this particular dating website/this particular psychology experiment]?
- How **comfortable** would you be referring a friend to [this particular clinic/this particular dating website/this particular psychology experiment]?

Composite Identity Threat for Study 1.

Based on a scree plot, these items showed some evidence of unidimensionality (Factor 1 accounted for 53% of total variance), so I created a composite score combining all 15 identity threat items into a single scale, which had high reliability ( $\alpha = 0.94$ ,  $M = 5.2$ ,  $SD = 1.3$ ).

## **Study 2.**

*Evaluation of Question.*

In Study 2, we asked four questions immediately following the experimentally assigned gender measure to assess participants' initial, general impressions. Participants indicated their answers on 7-point Likert scales (1 = Not at all, 7 = Very much).

- Overall, how much do you like this gender question?
- How fully and accurately does your answer reflect your OWN understanding of your gender?

- *How fully and accurately do you think OTHER PEOPLE would perceive your gender based on your answer?*
- *How much do you think OTHER PEOPLE will perceive your gender as real and legitimate based on your answer?*

These four questions had acceptable reliability ( $\alpha=.79$ ), so they were averaged into a single composite score ( $M= 5.1, SD = 1.3$ ).

*Authenticity (Kraus, Chen, and Keltner, 2011).*

We included a 4-item measure of felt authenticity developed by Kraus, Chen, and Keltner (2011). Participants reported how much they currently agreed with statements like “I feel like I can be myself with others” on a 7 point Likert scale (1 = disagree strongly, 7 = agree strongly). The items had acceptable reliability ( $\alpha = .76$ ) and were averaged ( $M = 5.1, SD = 1.2$ ).

*Collective Self-Esteem (Luhtanen & Crocker, 1992).*

We adapted the 16-item Collective Self-Esteem Scale developed by Luhtanen and Crocker (1992) to refer to “gender groups” instead of “social groups,” and used it with a 7 point Likert scale (1 = Strongly disagree, 7 = Strongly Agree). While Luhtanen and Crocker reported that the scale has four subscales (Membership, Private, Public, and Identity), principal components analysis of our data suggested that a 2-factor solution was more appropriate. Using a combination of our observed factor loadings and theoretical reasoning, we opted to create one subscale containing the items from Luhtanen and Crocker’s Membership/Private/Identity subscales (12 items,  $M=5.6, SD=0.94, \alpha=0.86$ ) and one subscale containing the Public items (4 items,  $M=4.8, SD=1.4, \alpha=0.79$ ). Typical items from the Membership/Private/Identity are “I am a worthy member of the social groups I belong to” and “In general, I’m glad to be a member of the social groups I belong to,” and a typical item from the Public scale is “In general, others respect the social groups that I am a member of.”

*Expectations of Belonging and Respect.*

Before answering these questions, participants saw a reminder of the gender demographic measure they answered at the beginning of the survey. They were asked to give their answers with respect to the hypothetical researchers who asked the questions at the beginning of the survey.

- *How likely do you think it is that they will treat you with respect?*
- *How likely do you think it is that they will reject you?*
- *How likely do you think it is that they will treat you like you belong?*
- *How likely do you think it is that they will treat you like an outsider?*
- *How likely do you think it is that they will treat you like you belong?*

Participants answered questions on a 7-point Likert scale (1 = Very unlikely, 7 = Very likely).

These questions were averaged into a single composite score ( $\alpha=.90$ ,  $M = 5.6$ ,  $SD = 1.4$ ).

#### *Comfort with/Investment in Research.*

Before answering these questions, participants saw a reminder of the gender demographic measure they answered at the beginning of the survey. They were asked to give their answers with respect to the hypothetical researchers who asked the questions at the beginning of the survey.

- *How comfortable would you be asking a friend to answer their survey/form?*
- *How willing would you be to recommend this survey to a friend?*
- *How comfortable would you feel continuing to answer their questions?*
- *How comfortable would you feel knowing that these people had this information about you?*
- *How comfortable would you feel completing the rest of their survey/form?*
- *How comfortable do you feel answering their gender question?*
- *How willing would you be to complete the rest of their survey/form?*

Participants answered questions on a 7-point Likert scale (1 = Not at all, 7 = Very much). These questions were averaged into a single composite score ( $\alpha=.93$ ,  $M = 5.4$ ,  $SD = 1.4$ ).

### **Perceptions of Researchers' Knowledge about Gender Minorities.**

#### *Study 1.*

Participants were asked to indicate their perceptions of the researchers who created these questions on several dimensions. Where not otherwise indicated, participants answered each question on a 7-point Likert scale (1 = Not at all, 7 = Very much).

*Researcher Knowledge about Gender.*

Participants answered each question on a 7-point Likert scale (1 = Almost nothing, 7 = A large amount). The composite score for these four items was reliable ( $\alpha = 0.95$ ,  $M = 4.7$ ,  $SD = 1.6$ ).

*How much do you think the people who created this gender question know about...*

- ...gender in general
- ...transgender people
- ...gender non-conforming people
- ...people with non-binary genders

*Question Consequences ( $\alpha = 0.94$ ,  $M = 4.5$ ,  $SD = 1.3$ ).*

Participants answered each question on a 7-point Likert scale (1 = Very negatively, 7 = Very positively).

*How do you think the use of this gender question will affect...*

- ...transgender people
- ...gender non-conforming people
- ...people with non-binary genders?
- ...you?

*Researcher Caring about Gender Minorities ( $\alpha = 0.97$ ,  $M = 5$ ,  $SD = 1.6$ ).*

*How much do you think the people who created this gender question care about the well-being of ...*

- ...transgender people
- ...gender non-conforming people
- ...people with non-binary genders

*Researcher Caring about You ( $\alpha = 0.88$ ,  $M = 4.5$ ,  $SD = 1.6$ ).*

*How much do you think the people who created this gender question...*

- ...care about your well-being
- ...understand you
- ...are interested in your opinions

### *Composite Perceptions of Researchers (Study 1).*

Based on a scree plot, these items showed some evidence of unidimensionality (Factor 1 accounted for 64% of total variance), so I created a composite score combining all 14 perceptions of researchers items into a single scale, which had high reliability ( $\alpha = 0.96$ ,  $M = 4.7$ ,  $SD = 1.3$ ).

### ***Study 2.***

Before answering each of these next questions, participants saw a reminder of the gender demographic measure they answered at the beginning of the survey. They were asked to give their answers with respect to the hypothetical researchers who asked the questions at the beginning of the survey. Unless otherwise indicated, participants answered questions on a 7-point Likert scale (1 = Very unlikely, 7 = Very likely):

#### *Perceptions of Researcher Knowledge/Concern about Transgender People.*

- *How likely do you think it is that they are concerned about non-binary people's well-being?*
- *How likely do you think it is that they consulted transgender people when creating the questions?*
- *How likely do you think it is that they are concerned about transgender people's well-being?*
- *How likely do you think it is that they consulted non-binary people when creating the questions?*
- *How likely do you think it is that they are transgender themselves?*

These questions had high reliability ( $\alpha=.93$ ), so they were averaged into a single composite score ( $M = 4.4$ ,  $SD = 1.7$ ).

### **Attitudes and Emotions toward Researchers.**

#### ***Attitudes toward Researchers (Study 1).***

Attitudes toward researchers were measured on five 7-point bipolar scales (1 = positive, 7 = negative) such as Good/Bad, Caring/Uncaring.

***Emotions Toward Researchers (Study 1 and Study 2).***

In both studies, participants were asked the degree to which they felt each of 10 emotions toward the researchers on a 7-point Likert scale (1=not at all, 7 = very much). Principal components analysis of our data suggested that a 2-factor solution was appropriate in both studies, so we created a subscale for positive emotions (Satisfaction, Empathy, Compassion, Pride) and a subscale for negative emotions (Fear, Irritation, Shame, Anxiety, Anger, Guilt). The positive emotions subscale (Study 1 M=3.5, SD=1.6; Study 2 M=3.3, SD=1.7) had good reliability (Study 1  $\alpha=0.86$ ; Study 2  $\alpha=.89$ ), as did the negative emotions subscale (Study 1 M=2.1, SD=1.3,  $\alpha=.91$ ; Study 2 M=2.1, SD=1.3,  $\alpha=.92$ ).

**Within-Participants Variables (Perceptions and Rankings of All Gender Measures).**

In the second half of each study, all participants answered each one of several different gender questions and provided feedback about each one. For more information about how these questions were selected and my predictions about each one, see Chapter 2. For more information about how responses to these questions were coded, see Chapter 3.

In Study 1, these were the Binary, Binary + Other, Multiple-Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Multiple-Choice + Transition History, Short Free Text, and Long Free Text questions. Four of these (Binary + Other, Multiple-Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity) were the same as those used in the experimental portion of Study 1, and the other four (Binary, Multiple-Choice + Transition History, Short Free Text, and Long Free Text) were new questions included for comparison purposes.

In Study 2, all participants answered all six gender questions that were used in the experimental portion of the study (Binary, Binary + Other, Multiple-Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Short Free Text).

### ***Perceptions of Gender Questions.***

Each gender measure was followed by questions about perceptions of that measure on 7-point Likert scales (1= Not at all, 7=Very much) and a free-text box for comments about that question/question pair.

In Study 1, these questions were:

- *How much did this gender question allow you to fully express your gender?*
- *How accurately does your answer to this gender question reflect your own understanding of your gender?*
- *How accurately do you think other people would perceive your gender based on your answer to this question?*
- *What comments do you have about this gender question?*

In Study 2, these questions were:

- *Overall, how much do you like [this gender question/these gender questions]?*
- *How fully and accurately [does your answer/do your answers] reflect your OWN understanding of your gender?*
- *How fully and accurately do you think OTHER PEOPLE would perceive your gender based on your [answer/answers]?*
- *How much do you think OTHER PEOPLE will perceive your gender as real and legitimate based on your [answer/answers]?*
- *What comments do you have about [this gender question/these two gender questions (as a pair)]?*

In each study, we combined the Likert-scale items about each gender measure into a single composite score for each of the measures. The scree plots for each of these 14 scales showed a clear, sharp drop after the first factor, supporting the hypothesis that these scales are unidimensional.

### ***Ranking of Gender Questions.***

After answering each of the different gender identity and transgender status measures and providing feedback about each one individually, participants ranked the measures compared to one another in order of their preference for each one (see Tables 4 and 6).

In Study 1, participants answered one question ranking the eight measures they had just seen, with the following prompt:

*You have now seen several different types of gender questions. Please rank these questions in order of your overall preference, with the question you like best in Position 1, and the question you like least in Position 8.*

In Study 2, participants were asked to rank all six measures they had just seen, but within a specific context. These contexts included the same three that were asked about experimentally in Study 1, as well as a new context – a job application. This new context was chosen because disclosure of gender identity and transgender status on a job application seems to carry the potential for discrimination, but not much potential benefit to participants (in contrast to clinics and dating websites, where disclosure may lead to better health care or better match suggestions). Participants answered the following question four times, about each of four different contexts (bracketed, in bold):

- *Imagine that you are completing a profile for a [**clinic intake form/psychology experiment/dating website/job application**]. Which of these questions would you most prefer to see in that context? Please rank these questions in order of your overall preference, with the question you like best in Position 1, and the question you like least in Position 7.*

The options were all previously answered six gender questions and a new option: “No gender question asked (Gender information is not collected at all).” This new option was included because gender information is routinely collected without a clear rationale or plan for the use of these data, and avoiding this sort of data collection would reduce the risks for participants inherent in disclosure. However, there is currently insufficient evidence about what positive value participants may place on this disclosure in different context, so we aimed to determine



how strongly participants valued having the opportunity to disclose their gender identity and transgender status in the first place.

## Results

### Between-Participants Analyses

Details about all significant results are reported in Appendix B.

#### Identity Threat.

##### *Study 1: Effects of Question, Context, and Transgender Identity.*

The three-way interaction of question (binary + other, multiple-choice, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity), context (clinic intake, dating website, or psychology experiment) and transgender identity (not transgender or transgender) had no significant effect on identity safety, perceptions of researchers, attitudes toward researchers, or emotions toward researchers. In those analyses, there were no two-way interactions or main effects of question for any of the dependent variables, so question was dropped from subsequent analyses.

However, there was a significant interaction of context and transgender identity for perceptions of researchers. Cisgender people in the dating website context perceived researchers as knowing and caring *less* about gender minorities, while transgender people in the dating context perceived researchers as knowing and caring *more* about gender minorities (Table B1). There were not enough unclassified participants for inclusion in this analysis.

There was a main effect of context on identity safety and perceptions of researchers' knowledge and caring (Table B2). Compared to participants in the clinic intake context, those in the dating website context reported less identity safety and perceived researchers as knowing and

caring less about gender minorities. Those in the psychology experiment context did not differ from those in the clinic intake context.

There was a main effect of transgender identity on identity safety, perceptions of researchers' knowledge and caring, attitudes toward researchers, and positive emotions toward researchers (Table B3). Compared to cisgender people, transgender people felt less identity safety and perceived researchers as knowing and caring less about gender minorities, but also had significantly more positive attitudes and emotions toward researchers. Unclassified participants did not differ from cisgender participants.

Table 1

*Results for Two-Way ANOVA Testing Effects of Context and Transgender Identity on Dependent Variables (Study 1).*

Dependent Variable	Context *		
	Trans Identity	Context	Trans Identity
Identity Safety	<i>ns</i>	F (2, 544) = 27.36, <i>p</i> <.001	F (2, 463) = 23.69, <i>p</i> <.001
Perceptions of Researchers' Knowledge/Caring	F (2, 443) = 3.70, <i>p</i> =.025	F (2, 520) = 4.89, <i>p</i> =.007	F (2, 461) = 8.54, <i>p</i> <.001
Attitudes toward Researchers	<i>ns</i>	<i>ns</i>	F (2, 447) = 4.71, <i>p</i> =.009
Positive Emotions toward Researchers	<i>ns</i>	<i>ns</i>	F (2, 459) = 3.49, <i>p</i> =.03
Negative Emotions toward Researchers	<i>ns</i>	<i>ns</i>	<i>ns</i>

*ns p* >.05

### ***Study 2: Effects of Question and Transgender Identity.***

The interaction of question (binary, binary + other, multiple-choice, multiple-choice + assigned sex at birth, multiple-choice + transgender identity, and short free text) and transgender identity (not transgender, transgender, or unclassified) had a significant effect on comfort with research. Compared to cisgender people, transgender people were significantly *less* comfortable with research when they saw the binary question, but significantly *more* comfortable when they saw any other question (Table B4). There was no significant interaction of question and transgender identity for any of the other dependent variables.

There was a main effect of question on several dependent variables (Table B5). Compared to those who saw the binary question, participants who saw the multiple-choice, multiple-choice + assigned sex at birth, multiple-choice + transgender identity, or short free text question had higher expectations of belonging, higher comfort with research, higher perceptions of researchers' knowledge of and caring about gender minorities, higher positive emotions toward researchers, and lower negative emotions toward researchers. Participants who saw the binary + other question also had higher perceptions of researchers' knowledge of and caring about gender minorities, and lower negative emotions toward researchers compared to those who saw the binary question, but did not differ from the those who saw the binary on any other variables. There were no significant effects of question on evaluation of question, authenticity, or CSES (public or private).

There was also a main effect of transgender identity on several dependent variables (Table B6). Compared to cisgender people, transgender participants and those who could not be classified had lower evaluation of questions, lower authenticity, lower CSES (both public and private), and lower expectations of belonging and respect. Unclassified participants had lower

comfort with research than cisgender people, but transgender people did not differ from cisgender people. Transgender people had more positive emotions toward researchers than cisgender people; unclassified people did not differ from cisgender people.

Table 2

*Two-Way ANOVAs Testing Effects of Question and Transgender Identity on Dependent Variables (Study 2)*

Dependent Variable	Question *		Trans Identity
	Trans Identity	Question	
Evaluation of Question	<i>ns</i>	<i>ns</i>	F (2, 583) = 24.72, <i>p</i> <.001
Authenticity	<i>ns</i>	<i>ns</i>	F (2, 588) = 32.43, <i>p</i> <.001
CSES (Private Subscale)	<i>ns</i>	<i>ns</i>	F (2, 570) = 26.18, <i>p</i> <.001
CSES (Public Subscale)	<i>ns</i>	<i>ns</i>	F (2, 580) = 52.28, <i>p</i> <.001
Expectation of Belonging/Respect	<i>ns</i>	F (5, 528) = 5.39, <i>p</i> <.001	F (2, 584) = 27.67, <i>p</i> <.001
Comfort with Research	F (10, 490) = 3.08, <i>p</i> <.001	F (5, 527) = 4.92, <i>p</i> <.001	F (2, 584) = 7.49, <i>p</i> <.001
Perceptions of Researchers	<i>ns</i>	F (5, 527) = 31.35, <i>p</i> <.001	<i>ns</i>
Positive Emotions Toward Researchers	<i>ns</i>	F (5, 522) = 9.30, <i>p</i> <.001	F (2, 579) = 4.81, <i>p</i> = .008
Negative Emotions Toward Researchers	<i>ns</i>	F (5, 523) = 4.07, <i>p</i> = .001	<i>ns</i>

*ns p* >.05

## **Within-Participants Analyses**

Details about all significant t-tests are reported in Appendix B.

### **Study 1.**

#### ***Question perceptions.***

Across all participants, the binary question was perceived as the least accurate and complete, and the binary + other and multiple-choice + transition history had the next lowest evaluations. The multiple-choice, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity questions were perceived as more accurate and complete, and the short and long free-text questions were perceived as most accurate and complete (Table 3).

The interaction of gender identity (man, woman, or non-binary/agender) and transgender identity (not transgender, transgender, or unclassified) did not have a significant effect on perceptions of any question.

There was a main effect of gender identity on perception of every question except long free-text (Table B7). Non-binary/agender people perceived all questions as less accurate and complete than men did. Women perceived the binary question as less accurate and complete than men did, but did not differ from men on their perceptions of any other question.

There was a main effect of transgender identity on perception of every question (Table B8). Compared to non-transgender people, transgender people perceived every question except long free-text as less accurate and complete, and unclassified people perceived all questions as less accurate and complete.

#### ***Question rankings.***

The interaction of gender identity (man, woman, or non-binary/agender) and transgender identity (not transgender, transgender, or unclassified) had a significant effect on rankings of the

multiple-choice + assigned sex at birth, short free-text, and long free-text questions (Table B9). For the multiple-choice + assigned sex at birth question, women and non-binary/agender people who were unclassified on trans identity assigned the question a lower ranking. For the short free-text question, transgender women assigned the question a lower ranking, but unclassified women assigned the question a higher ranking. For the long free-text question, women and non-binary/agender people who were unclassified on trans identity assigned the question a higher ranking.

There was a main effect of gender identity on rankings of all questions except multiple-choice + transition history. Compared to men, non-binary/agender people assigned lower rankings to the binary, binary + other, and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice, multiple-choice + transgender identity, short free-text, and long free-text questions. Women assigned lower rankings to the binary question compared to men, but did not differ from men in their rankings of any other question.

There was also a main effect of transgender identity for on rankings of all questions except multiple-choice + transition history. Compared to cisgender people, transgender people assigned lower rankings to the binary, binary + other, and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice, multiple-choice + transgender identity, short free-text, and long free-text questions. Unclassified people assigned lower rankings to the binary question compared to cisgender people, but did not differ from cisgender people in their rankings of any other question.



Table 3

*Descriptive Statistics for Composite Scores Assessing Perceptions of Gender Measures (Studies 1 and 2)*

Gender Measure Evaluated		M	SD	$\alpha$
Study 1 (3 items)	Binary	4	2.3	0.93
	Binary + Other	5.4	1.6	0.86
	Multiple-Choice	5.7	1.5	0.85
	Multiple-Choice + Assigned Sex at Birth	5.6	1.5	0.88
	Multiple-Choice + Transgender Identity	5.7	1.5	0.89
	Multiple-Choice + Transition History	5.3	1.7	0.92
	Short Free-Text	6.1	1.2	0.85
	Long Free-Text	6.4	1.1	0.87
Study 2 (4 items)	Binary	4.9	1.7	0.85
	Binary +Other	5.3	1.5	0.85
	Multiple-Choice	5.6	1.3	0.84
	Multiple-Choice + Assigned Sex at Birth	5.4	1.4	0.83
	Multiple-Choice + Transgender Identity	5.4	1.4	0.85
	Short Free-Text	5.8	1.4	0.88

Table 4  
*Rankings of Gender Questions (Study 1)*

Gender Question	Mean Ranking	SD	Median Ranking
Short Free-Text	3.41	2.26	2
Long Free-Text	3.47	2.76	2
Multiple-Choice	3.75	1.57	4
Multiple-Choice + Transgender Identity	4.17	1.58	4
Multiple-Choice + Assigned Sex at Birth	4.60	1.85	5
Binary + Other	4.75	2.07	5
Multiple-Choice + Transition History	5.42	1.63	6
Binary	6.43	2.54	8

*Note.* Higher numbers indicate lower ranking (i.e. 8 is the lowest possible ranking, and 1 is the highest).

Table 5

*Two-Way ANOVAs Testing Effects of Gender Identity and Transgender Identity on Perceptions of Questions (Study 1)*

Dependent Variable	Gender Identity * Trans Identity	Gender Identity	Trans Identity
Binary	<i>ns</i>	F (2, 422) = 256.65, <i>p</i> <.001	F (2, 475) = 208.46, <i>p</i> <.001
Binary + Other	<i>ns</i>	F (2, 421) = 16.35, <i>p</i> <.001	F (2, 472) = 49.545, <i>p</i> <.001
Multiple-Choice	<i>ns</i>	F (2, 422) = 4.44, <i>p</i> =.012	F (2, 473) = 32.04, <i>p</i> <.001
MC + Assigned Sex at Birth	<i>ns</i>	F (2, 421) = 20.61, <i>p</i> <.001	F (2, 472) = 71.06, <i>p</i> <.001
MC + Transgender Identity	<i>ns</i>	F (2, 422) = 5.83, <i>p</i> =.003	F (2, 474) = 45.33, <i>p</i> <.001
MC + Transition History	<i>ns</i>	F (2, 421) = 29.21, <i>p</i> <.001	F (2, 469) = 73.34, <i>p</i> <.001
Short Free-Text	<i>ns</i>	F (2, 418) = 8.23, <i>p</i> <.001	F (2, 467) = 18.18, <i>p</i> <.001
Long Free-Text	<i>ns</i>	<i>ns</i>	F (2, 464) = 12.42, <i>p</i> <.001

*ns p* >.05

Table 6

*Two-Way ANOVAs Testing Effects of Gender Identity and Transgender Identity on Rankings of Questions (Study 1)*

Dependent Variable	Gender Identity *		
	Trans Identity	Gender Identity	Trans Identity
Binary	<i>ns</i>	F (2, 373) = 37.18, <i>p</i> <.001	F (2, 421) = 47.85, <i>p</i> <.001
Binary + Other	<i>ns</i>	F (2, 373) = 11.24, <i>p</i> <.001	F (2, 421) = 15.79, <i>p</i> <.001
Multiple-Choice	<i>ns</i>	F (2, 373) = 4.89, <i>p</i> =.008	F (2, 421) = 3.20, <i>p</i> =.04
MC + Assigned Sex at Birth	F (4, 366) = 3.58, <i>p</i> = .007	F (2, 373) = 3.39, <i>p</i> =.03	F (2, 421) = 11.36, <i>p</i> <.001
MC + Transgender Identity	<i>ns</i>	F (2, 373) = 10.55, <i>p</i> <.001	F (2, 421) = 17.30, <i>p</i> <.001
MC + Transition History	<i>ns</i>	<i>ns</i>	<i>ns</i>
Short Free-Text	F (4, 366) = 3.84, <i>p</i> = .004	F (2, 373) = 6.74, <i>p</i> <.001	F (2, 421) = 18.09, <i>p</i> <.001
Long Free-Text	F (4, 366) = 4.13, <i>p</i> = .002	F (2, 373) = 13.58, <i>p</i> <.001	F (2, 421) = 18.08, <i>p</i> <.001

*Note.* Higher numbers indicate lower ranking (i.e. 8 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

## **Study 2.**

### ***Question perceptions.***

Across all participants, the binary question was perceived as the least accurate and complete. The binary + other, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity questions were perceived as more accurate and complete, and the multiple-choice and short free-text questions were perceived as most accurate and complete (Table 3).

The interaction of gender identity (man, woman, or non-binary/agender) and transgender identity (not transgender, transgender, or unclassified) did not have a significant effect on perceptions of any question (Table 7). However, there were main effects of both gender identity and transgender identity on perceptions of all questions (Table B12 and Table B13).

Compared to men, non-binary/agender people perceived all questions as *less* accurate and complete. Women evaluated the multiple-choice, multiple-choice + assigned sex at birth, multiple-choice + transgender identity, and short free-text questions as *more* accurate and complete than men did, but evaluated the binary question as *less* accurate and complete. Women did not differ from men in their perceptions of the binary + other question.

Compared to cisgender people, transgender people perceived all gender questions as less accurate and complete. Unclassified people also perceived all questions except multiple-choice + assigned sex at birth as less accurate and complete than cisgender people did.

Table 7

*Two-Way ANOVAs Testing Effects of Gender Identity and Transgender Identity on Perceptions of Questions (Study 2)*

Question Type	Gender Identity *		
	Trans Identity	Gender Identity	Trans Identity
Binary	<i>ns</i>	F (2, 572) = 40.51, <i>p</i> < .001	F (2, 583) = 33.27, <i>p</i> < .001
Binary + Other	<i>ns</i>	F (2, 568) = 48.05, <i>p</i> < .001	F (2, 578) = 38.68, <i>p</i> < .001
Multiple-Choice	<i>ns</i>	F (2, 574) = 23.83, <i>p</i> < .001	F (2, 583) = 16.14, <i>p</i> < .001
MC + Assigned Sex at Birth	<i>ns</i>	F (2, 566) = 38.93, <i>p</i> < .001	F (2, 575) = 40.09, <i>p</i> < .001
MC + Transgender Identity	<i>ns</i>	F (2, 576) = 14.60, <i>p</i> < .001	F (2, 588) = 15.81, <i>p</i> < .001
Short Free-Text	<i>ns</i>	F (2, 573) = 20.53, <i>p</i> < .001	F (2, 583) = 24.89, <i>p</i> < .001

*ns p* > .05

### *Question rankings.*

In the clinic intake context and psychology experiment context, the binary and no question asked options were ranked the lowest, followed by the binary + other question. The next highest ranked options were the multiple-choice, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity. The short free-text question was ranked most highly.

In the dating website context, the binary and no question asked options were ranked the lowest. The next highest ranked options were the binary + other, multiple-choice, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity. The short free-text question was ranked most highly.

In the job application context, the binary, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity were ranked the lowest. The next highest ranked options were the binary + other and the multiple-choice question. The short free-text question and no question asked options were ranked most highly.

The interaction of gender identity and transgender identity did not have a significant effect on rankings of any question in the psychology experiment, dating website, or job application contexts. In the clinic intake context only, the interaction had a significant effect on rankings of the multiple-choice + transgender identity question. Women who identified as transgender ( $M = 5.78$ ,  $SE = 0.73$ ,  $diff. = 1.66$ ) ranked this question less highly ( $t(4, 472) = 2.26$ ,  $p = .023$ ). Similarly, non-binary/agender people who identified as transgender ( $M = 6.64$ ,  $SE = 0.81$ ,  $diff. = 2.52$ ) ranked this question less highly ( $t(4, 472) = 3.09$ ,  $p = .002$ ).

### *Main effects of gender identity on question rankings by context.*

In the clinic intake and psychology experiment contexts, there was a main effect of gender identity for rankings of the binary, binary + other, multiple-choice, and multiple-choice +

transgender identity questions (Table B14 and B15). In both contexts, women and nonbinary/agender people gave lower rankings to the binary and binary + other questions, and higher rankings to the multiple-choice and multiple-choice + transgender identity questions than men did.

In the dating website context, there was a main effect of gender identity for rankings of every question (Table B16). Women gave lower rankings to the binary and binary + other questions, and higher rankings to the multiple-choice, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity than men did. Women and men did not differ in their rankings of the no question asked option. Non-binary/agender people gave lower rankings to the binary, binary + other, and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice, multiple-choice + transgender identity, short free-text, and no question asked option than men did.

In the job application context, there was a main effect of gender identity for rankings of every question except multiple-choice + transgender identity (Table B17). Women gave lower rankings to the binary and binary + other questions, and higher rankings to the multiple-choice, short free-text, and no question asked option than men did. Women and men did not differ in their rankings of the multiple-choice + assigned sex at birth question. Non-binary/agender people gave lower rankings to the binary, binary + other, and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice, short free-text, and no question asked option than men did.

*Main effects of transgender identity on question rankings.*

In the clinic intake context, there was a main effect of transgender identity for rankings of the binary, binary + other, multiple-choice, and multiple-choice + transgender identity questions

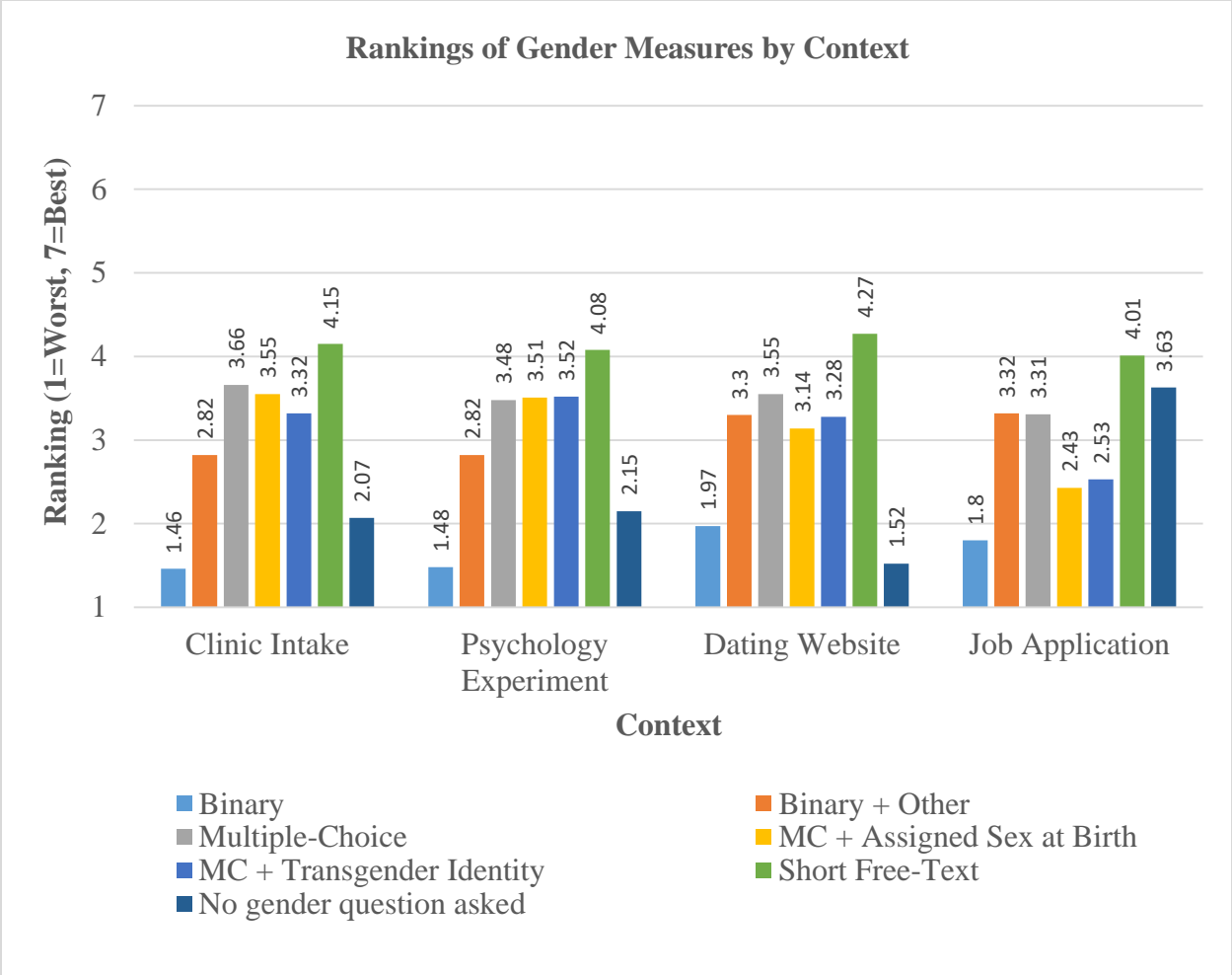


(Table B18). Transgender people gave lower rankings to the binary and binary + other questions, and higher rankings to the multiple-choice and multiple-choice + transgender identity questions, than cisgender people did. Unclassified people gave lower rankings to the binary + other question, and higher rankings to the multiple choice question than cisgender people did, but otherwise did not differ from cisgender people in their rankings.

In the psychology experiment context, there was a main effect of transgender identity for rankings of the binary and multiple-choice + transgender identity questions (Table B19). Transgender people gave lower rankings to the binary question and higher rankings to the multiple-choice + transgender identity question than cisgender people did. Unclassified people did not differ in their rankings from cisgender people.

In the dating website context, there was a main effect of transgender identity for rankings of the binary, multiple-choice, multiple-choice + assigned sex at birth, short free-text, and no question asked questions (Table B20). Transgender people gave lower rankings to the binary and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice, short free-text, and no question asked questions than cisgender people did. Unclassified people gave lower rankings to the binary question than cisgender people did, but otherwise did not differ in their rankings from cisgender people.

In the job application context, there was a main effect of transgender identity for rankings of the binary, multiple-choice, multiple-choice + assigned sex at birth, and no question asked questions (Table B21). Transgender people gave lower rankings to the binary and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice and no question asked questions than cisgender people did. Unclassified people did not differ in their rankings from cisgender people.



*Figure 1. Rankings of Gender Measures by Context (Study 2).*  
 Note. Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

Table 8

*Two-Way ANOVAs Testing Effects of Gender Identity and Transgender Identity on Rankings of Questions (Study 2)*

Question	Gender Identity *			
	Trans Identity	Gender Identity	Trans Identity	
Clinic Intake Context	Binary	<i>ns</i>	F (2,482) = 24.01, <i>p</i> <.001	F (2,490) = 8.31, <i>p</i> <.001
	Binary + Other	<i>ns</i>	F (2,482) = 17.33, <i>p</i> <.001	F (2,490) = 6.15, <i>p</i> =.002
	Multiple-Choice	<i>ns</i>	F (2,482) = 6.62, <i>p</i> =.001	F (2,490) = 5.21, <i>p</i> =.005
	MC + Assigned Sex at Birth	<i>ns</i>	<i>ns</i>	<i>ns</i>
	MC + Transgender Identity	F (4,472) = 2.71, <i>p</i> =.02	F (2,482) = 11.41, <i>p</i> <.001	F (2,490) = 5.88, <i>p</i> =.002
	Short Free-Text	<i>ns</i>	<i>ns</i>	<i>ns</i>
	No Question Asked	<i>ns</i>	<i>ns</i>	<i>ns</i>
Psychology Context	Binary	<i>ns</i>	F (2,457) = 21.99, <i>p</i> <.001	F (2,464) = 9.00, <i>p</i> =.002
	Binary + Other	<i>ns</i>	F (2,457) = 13.75, <i>p</i> <.001	<i>ns</i>
	Multiple-Choice	<i>ns</i>	F (2,457) = 10.27, <i>p</i> <.001	<i>ns</i>
	MC + Assigned Sex at Birth	<i>ns</i>	<i>ns</i>	<i>ns</i>
	MC + Transgender Identity	<i>ns</i>	F (2,457) = 9.13, <i>p</i> <.001	F (2,464) = 4.42, <i>p</i> =.01
	Short Free-Text	<i>ns</i>	<i>ns</i>	<i>ns</i>
	No Question Asked	<i>ns</i>	<i>ns</i>	<i>ns</i>
Dating Website Context	Binary	<i>ns</i>	F (2,468) = 29.98, <i>p</i> <.001	F (2,475) = 12.84, <i>p</i> <.001
	Binary + Other	<i>ns</i>	F (2,468) = 9.90, <i>p</i> <.001	<i>ns</i>
	Multiple-Choice	<i>ns</i>	F (2,468) = 16.64, <i>p</i> <.001	F (2,475) = 12.24, <i>p</i> <.001
	MC + Assigned Sex at Birth	<i>ns</i>	F (2,468) = 12.60, <i>p</i> <.001	F (2,475) = 8.23, <i>p</i> <.001
	MC + Transgender Identity	<i>ns</i>	F (2,468) = 3.50, <i>p</i> =.03	<i>ns</i>
	Short Free-Text	<i>ns</i>	F (2,468) = 11.71, <i>p</i> <.001	F (2,475) = 3.42, <i>p</i> =.03

	No Question Asked	<i>ns</i>	F (2,468) = 7.52, <i>p</i> <.001	F (2,475) = 4.81, <i>p</i> =.008
Job Application Context	Binary	<i>ns</i>	F (2,460) = 14.93, <i>p</i> <.001	F (2,468) = 10.04, <i>p</i> <.001
	Binary + Other	<i>ns</i>	F (2,460) = 7.37, <i>p</i> <.001	<i>ns</i>
	Multiple-Choice	<i>ns</i>	F (2,460) = 8.23, <i>p</i> <.001	F (2,468) = 6.17, <i>p</i> =.002
	MC + Assigned Sex at Birth	<i>ns</i>	F (2,460) = 9.60, <i>p</i> <.001	F (2,468) = 10.52, <i>p</i> <.001
	MC + Transgender Identity	<i>ns</i>	<i>ns</i>	<i>ns</i>
	Short Free-Text	<i>ns</i>	F (2,460) = 7.66, <i>p</i> <.001	<i>ns</i>
	No Question Asked	<i>ns</i>	F (2,460) = 13.33, <i>p</i> <.001	F (2,468) = 10.61, <i>p</i> <.001
	<i>ns p</i> >.05			

## Discussion

### Do Gender Questions Cue Identity Safety and/or Identity Denial?

These results suggest that wording of gender questions does serve as a cue to identity safety for transgender participants, and that seeing inclusive questions improves views of researchers and comfort with research. While there was no significant effect of question on any of the dependent variables for Study 1, all experimentally tested questions in Study 1 were at least nominally inclusive of transgender people. In Study 2, with the addition of the binary question as a baseline, there was a significant interaction of question and transgender identity for comfort with research. Transgender people were significantly *less* comfortable with research when they saw the binary question, and significantly *more* comfortable when they saw any other question, suggesting that the use of inclusive gender questions does serve as a cue to identity safety, increasing comfort with research and likely increasing rates of participation among transgender people.

This increase in comfort is particularly important because the baseline for transgender participants was lower than for cisgender ones. Across questions, transgender participants felt less authentic, had lower public and private collective self-esteem, and had lower expectations of belonging and respect than cisgender participants. While there was no direct effect of question on collective self-esteem or authenticity, our measures of those constructs ask about general experiences, while the effect of a gender question seems likely to be quite localized to a situation. Transgender people also evaluated questions less positively across the board, perhaps suggesting that they are more sensitive to the wording of any given question, and thus less satisfied with any specific option. These results are consistent with the idea that transgender

people experience generally lower identity safety and feel less positively about their identities than cisgender people.

Across participants, those who saw any question other than the binary perceived researchers as knowing and caring more about gender minorities, and had fewer negative emotions toward researchers. However, not all inclusive questions convey the same levels of identity safety. The binary + other question did not increase feelings of belonging, comfort with research, or positive emotions toward researchers relative to the binary question (presumably because of the identity denial they pose for non-binary and agender participants), but all other inclusive questions did. Those who saw any question other than the binary or binary + other also had higher expectations of belonging, higher comfort with research, and more positive emotions toward researchers. This pattern of results suggests that both cisgender and transgender participants notice the presence of inclusive questions, and use this to infer researchers' views of gender minorities. Seeing inclusive gender questions improved perceptions of researchers and made both cisgender and transgender participants more comfortable with research. Thus, using inclusive gender questions in research may increase retention of both transgender participants and cisgender ones.

### **What Effects Does Context Have on Perceptions of Gender Questions?**

#### **Study 1: Between-Participants Context Effects**

Study 1 included context as an experimental variable. Overall, participants found the dating website to be less identity safe than the clinic intake or psychology experiment conditions, and perceived researchers in that condition as less caring and less informed about gender minorities. This is consistent with our previous findings that the salience of gender and motivation to be perceived accurately is heightened in dating contexts relative to clinic intake

contexts among cisgender people (Hauptert, Murphy, & Smith, 2014). It seems that the way in which gender information is collected on dating websites acquires greater significance than in other contexts, presumably because information about gender is an essential component of dating for most people.

In addition to the main effect of context, there was a significant interaction of context and transgender identity on perceptions of researcher caring; cisgender people in the dating website context perceived researchers as knowing and caring *less* about gender minorities, while transgender people in the dating context perceived researchers as knowing and caring *more* about gender minorities. This appears to be an area of misperception between cisgender and transgender people; cisgender people perceived these questions as relatively uninformed and uncaring towards transgender people, while transgender people themselves perceived the questions as relatively more inclusive and informed. Because all experimentally tested questions in Study 1 were at least nominally inclusive of transgender people, we interpret this to mean that transgender people value the opportunity to disclose their transgender status in a dating context, and thus perceive an inclusive question in that context particularly positively. In contrast, cisgender people seem to misperceive how transgender people will feel about disclosing their transgender status, and undervalue the opportunity to do this.

The free-text comments about what participants disliked about each question also provide some support for this interpretation. For example, one cisgender woman (undergraduate student sample, dating context, multiple-choice + sex assigned at birth question) commented “Although you gave many options it may be controversial because trans individuals may not be comfortable with stating that on paperwork so they will still just choose male or female.” In contrast, a

participant who identified as “CAMAB<sup>17</sup> genderqueer woman” (gender minority sample, dating context, multiple-choice question) commented “I was not allowed to select multiple boxes (e.g. genderqueer woman); I was not allowed to identify that I was transgender.” Another participant who identified as “Trans woman, currently unable to transition” (gender minority sample, dating context, binary + other question) commented

“Why would questions of preoccupation and focus on gender crop up on a dating website?<sup>18</sup> There is additionally no distinguishment between trans and cis identities. I would consider accounting for trans status to be an important and confidential feature of any dating site that respects gender minorities.”

Taken together, this anecdotal evidence supports the quantitative findings and suggests that having some opportunity to disclose transgender status in a dating context is generally valuable to transgender people, and likely more valuable than cisgender people anticipate.

### **Study 2: Within-Participants Context Effects**

In Study 2, we included context as a within-participants variable rather than an experimental variable, and added a job application context alongside the clinic intake, psychological experiment, and dating website contexts used in Study 1. Across contexts and participants, the binary question was ranked lowest, and the short free-text question was ranked highest, consistent with our findings in Study 1. This is unsurprising given the many objectionable features of the binary question (see Chapter 1) and the fully customizable nature of the short free-text question. As this was consistent across all groups and conditions, we do not discuss this finding further.

### **Clinic intake and psychology experiment contexts.**

---

<sup>17</sup> Coercively Assigned Male At Birth

<sup>18</sup> This is a reference to the questions about how much you are currently thinking about/focusing on your gender, which some participants believed were intended as part of the gender question itself in Study 1.



The pattern of results in the clinic intake and psychology experiment contexts was quite similar. In both contexts, the multiple-choice, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity questions were ranked about the same (and second only to the short free-text). The binary + other question was ranked lower, and the no question asked option lower still. For both contexts, transgender participants ranked the binary question lower and the multiple-choice + transgender identity question higher than cisgender participants, and in the clinic context, transgender participants also ranked the binary + other question lower and the multiple-choice question higher than cisgender participants. Likewise, in both contexts women and non-binary/agender participants ranked the binary and binary + other questions lower and the multiple-choice and multiple-choice + transgender identity questions higher than men.

These results suggest that in clinic and psychological contexts, participants generally want to disclose more information about their identities (as evidenced by the low rankings of no question asked, binary, and binary + other questions). The questions about transgender status also seem broadly acceptable to participants, with a question about transgender identity being particularly well-received by transgender participants. It may be that disclosing transgender status is perceived as lower-risk (or at least as having a better ratio of benefits to costs) in clinical and psychological contexts relative to other contexts, which may increase participants' desire to disclose more detailed information about their identities.

#### **Dating website context.**

In the dating website context, the binary + other, multiple-choice, multiple-choice + sex assigned at birth, and multiple-choice + transgender identity questions were all ranked roughly equally, with no question asked and the binary question ranking lowest. Indeed, the binary + other question was regarded more positively in the dating context than in any other context. In

the dating context, transgender people ranked the multiple-choice, short free-text, and no question asked questions more highly, and the binary and multiple-choice + assigned sex at birth question lower than did cisgender people. Likewise, non-binary/agender people ranked the multiple-choice, short free-text, and no question asked questions more highly, and the binary, binary + other, and multiple-choice + assigned sex at birth question lower than did men. However, cisgender and transgender people did not differ in their rankings of the binary + other and multiple-choice + transgender identity questions.

In combination with our findings from Study 1, these results suggest that many transgender people value the opportunity to disclose their transgender status in dating contexts, but that the type of question matters. Transgender people and non-binary/agender people preferred the multiple-choice + transgender identity question over the multiple-choice + assigned sex at birth question, presumably because the former emphasizes self-definition and identity instead of the physical body (see Chapter 2 for a full discussion). However, the higher ranking of the multiple-choice and no question asked options among transgender people relative to cisgender people also suggests that there are some transgender people who would prefer not to disclose their transgender status in dating contexts. This tension may reflect the fact that disclosing transgender status in a dating context carries a greater risk than in other contexts (i.e. of intimate partner violence; Dank et al., 2014; James et al., 2016) but also the possible benefit of avoiding dates with transphobic partners. Different people will weigh these risks and benefits differently, which in turn will influence their question preferences. For general use, we feel that these results most support the use of the multiple-choice + transgender identity question (which was highly ranked by both cisgender and transgender people of all genders), but they also

indicate the importance of allowing participants to opt out of answering the question by including a “prefer not to answer” option and/or making the entire question optional.

### **Job application context.**

In the job application context, the short free-text question and no question asked options were rated most highly, followed by the binary + other and multiple-choice questions. The binary, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity were ranked lowest. Transgender people gave lower rankings to the binary and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice and no question asked questions than cisgender people did. Similarly, non-binary/agender people gave lower rankings to the binary, binary + other, and multiple-choice + assigned sex at birth questions, and higher rankings to the multiple-choice, short free-text, and no question asked option than men did.

The most highly-ranked questions in the job application context (short free-text and no question asked) permit participants to disclose the smallest amount of information about themselves, and the next most highly-ranked questions (binary + other and multiple-choice) allow expression of gender identity without asking about transgender identity in any way. Disclosing gender information, and especially transgender status, has the potential to result in employment discrimination, but very little potential benefit. Indeed, a job application that collects gender information poses sufficient risk of discrimination in most cases that an applicant who encounters one can file a sex discrimination lawsuit (EEOC, 2018). In this context, obviously, we would suggest not collecting gender information at all (at least not as part of the application itself).

### **Which questions do participants prefer?**

As expected, across both studies, the short free-text question was perceived as most accurate and complete, and ranked highest regardless of participant gender, transgender status, or context. It was closely followed by the multiple-choice question, and the binary + other, multiple-choice + assigned sex at birth, and multiple-choice + transgender identity questions were next most liked. The binary question was perceived as least accurate and complete, and ranked lowest regardless of participant gender, transgender status, or context in both studies.

Preferences did differ between participants. In Study 2, overall, non-binary/agender people gave lower rankings to the binary and binary + other questions, and higher rankings to the multiple-choice and multiple-choice + transgender identity questions compared to men. Likewise, across contexts, transgender people gave lower rankings to the binary question than cisgender people did. This further strengthens our case that the binary and binary + other questions are not good choices for researchers in any context. The frequent similarity in response patterns between women and non-binary participants relative to men (e.g. women, like non-binary people, evaluated the multiple-choice, multiple-choice + assigned sex at birth, multiple-choice + transgender identity, and short free-text questions as *more* accurate and complete than men did, but evaluated the binary question as *less* accurate and complete) may also suggest opportunities for coalition-building among women and non-binary/agender people; both groups face discrimination and prejudice on the basis of gender, and have more reason than men to be concerned by the way researchers collect and use this information.

The two measures of transgender status (multiple-choice + assigned sex at birth and multiple-choice + transgender identity) were usually ranked comparably by cisgender people, but transgender people and non-binary/agender people preferred the multiple-choice + transgender identity question in most contexts. As we suggested in Chapter 2, this makes sense given that a

question about transgender identity empowers participants to describe themselves as they choose, while a question about assigned sex at birth privileges the physical body and leaves attributions about its meaning to the researcher.

### **Which questions should researchers use?**

As discussed in previous chapters, we do not recommend the binary question for use in any context because of its many theoretical issues. This chapter also presents evidence that the binary question poses a risk of identity threat for transgender and non-binary people. For different reasons, we do not recommend the short free-text question. While it is very well-liked and creates a sense of identity safety, coding responses is challenging and time-consuming, and ultimately removes agency from participants and places it on researchers (see Chapter 3). Therefore, we do not recommend either the binary question or the short free-text question.

When information about transgender status is not needed, these results provide the strongest support for use of the multiple-choice question. It was very well-liked across participants and contexts in both studies, and appeared to create a greater sense of identity safety among participants than the binary or binary + other questions. In Study 2, participants who were experimentally assigned to see the multiple-choice question had higher expectations of belonging, higher comfort with research, and lower negative emotions toward researchers than those who saw the binary. In contrast, those who saw the binary + other question did not differ from those who saw the binary question on those measures.

When information about transgender status is necessary, we suggest using the multiple-choice + transgender identity question. The rankings of the multiple-choice + transition history question in Study 1 were low enough that we did not include it in Study 2 at all. The multiple-choice + assigned sex at birth and multiple-choice + transgender identity questions did not differ

in the identity safety they cued, and both were somewhat less well-liked than the multiple-choice question alone. However, while cisgender people and men ranked the two questions similarly, transgender people and non-binary/agender people ranked the multiple-choice + transgender identity question more highly in most contexts. Given that the multiple-choice + transgender identity question also yields more readily-usable results (especially for non-binary/agender people) than the multiple-choice + assigned sex at birth question (as shown in Chapter 3), we recommend the multiple-choice + transgender identity question for general use. That said, the multiple-choice + assigned sex at birth question also performed well, and may be still be useful, particularly in clinical contexts.

Alongside all of these questions, researchers should explain their considerations in choosing this particular question (e.g. number of groups for statistical analyses), and what exactly they will use the information for (e.g. Will it be used simply to ensure a diverse participant pool, or will responses between groups be compared? Are researchers more interested in bodies or social perceptions?). As described in Chapter 1, this information helps participants give researchers the information they are seeking, and reduces the potential for miscommunication. It may even serve as another cue to identity safety, helping gender minority participants to see that researchers have considered their situation and care about their responses.

### **Conclusion**

We recommend the multiple-choice and multiple-choice + transgender identity questions for general use, as they were identity safe and well-liked across contexts, particularly for transgender and non-binary/agender participants. They also meet our other considerations for simplicity of collection and analysis, as discussed in Chapter 3. In Chapter 5, we present

evidence for the validity of the multiple-choice question (tested with and without the transgender identity question).

## **Chapter 5: Validating Inclusive Gender Identity Measures**

In Chapter 3, I discussed the process of data cleaning for each of several different gender identity and transgender status measures, and recommended the multiple-choice gender identity measure and transgender identity measure as having high sensitivity, content validity, and generalizability while remaining comparatively simple to administer and analyze. In Chapter 4, I presented experimental evidence about the risk of identity threat posed by each of these gender identity and transgender status measures, and identified the multiple-choice gender identity measure (with or without a transgender identity measure) as the least threatening across contexts and most liked by a wide range of participants (that is, with the best consequential validity). In this chapter, I present evidence from two studies which examined the substantive validity, generalizability, and external validity of the multiple-choice gender identity measure, alongside the transgender identity measure.

### **Types of Validity**

To achieve acceptance in psychological research, the proposed gender identity measure needs to be theory-grounded and rigorously validated. Researchers are unlikely to replace the current, familiar binary gender question with the proposed gender identity measure unless it has significant advantages in terms of precision or is much better at predicting outcomes of interest. The next sections describe several different types of validity and the evidence I will present about each type.

#### **Substantive Validity**

Substantive validity is typically established by using a measure to distinguish criterion and contrast groups, to predict behavior over time, or to demonstrate prediction or moderation of another process (John & Benet-Martinez, 2000). I argue that gender identity is conceptually



most related to the way a person views the relationship between themselves and a given gender group, and thus need not perfectly predict gendered appearance, stereotypical behavior, or other more distant gendered constructs. Likewise, while a measure's substantive validity is frequently established through its ability to distinguish criterion and contrast groups, in the current situation criterion and contrast groups are themselves already distinguished by imperfect methods. The general human tendency is to immediately and uncontrollably perceive individuals as either men or women, and subsequently to regard these perceptions as objective even in the face of substantial contradictory evidence and/or conflicting perceptions from other people (e.g. Kessler & McKenna, 1978). For this reason, there may be some discrepancies between results on this measure and gender criteria (e.g. physical appearance, sex on passport) which are not attributable to a lack of substantive validity. Nonetheless, there are several behaviors and processes which are clearly related exclusively or primarily to identity that may be useful for demonstrating the substantive validity of the multiple-choice gender identity question.

#### **Evidence from self-definition and self-investment.**

Leach et al. (2008) subdivide group identity into *self-definition* (a subjective assessment of the degree of similarity between the members of a given gender group and the self, and homogeneity of the available gender groups in the culture), and *self-investment* (a subjective assessment of the centrality of a given gender group to one's sense of self and of the satisfaction and solidarity one feels in relation to that gender group). In Study 1, I test the ability of each gender identity question to predict the levels of self-definition and self-investment participants feel with the groups "men" and "women" using the Leach et al. (2008) Multi-Dimensional Social Identity Scale. For example, if men show higher self-definition and self-investment in the group "men" than people of other genders, this will provide evidence of substantive validity. I also

measured identification with each of several different gender groups using versions of a single-item social identification scale designed by Postmes et al. (2013).

### **Evidence from identity denial.**

The experience of *identity denial* should also theoretically be influenced by gender identity. McLemore (2014) has shown that transgender people who experience identity denial in the form of misgendering (i.e. being treated by others as a gender other than the one with which they identify) feel stigmatized and less authentic as a result. Generalizing from this finding, we can presume that participants who are described by pronouns inconsistent with their gender identities (e.g. a woman being called “he,” a man being called “she,” or a genderqueer person being called “he” or “she”) will feel discomfort. However, participants who are described with identity-consistent pronouns will not experience identity denial and thus should not report discomfort. Similarly, we expect that participants who are required to select identity-denying gender options on official forms will experience discomfort. Many transgender people must do this in order to access needed resources (e.g. if a person’s sex on health insurance forms does not match their claim, the bill may be denied). In Study 2, I used five items adapted from the Multi-Gender Identity Questionnaire (Joel et al., 2014) to measure this feeling of identity denial in relation to others’ pronoun use and the collection of gender information on administrative forms (e.g. *In the past 12 months, have you been upset when other people used female pronouns (she/her) to refer to you?*).

### **Generalizability**

I provided evidence about the reliability component of the generalizability of the multiple-choice gender identity question in Chapter 3 (i.e. test-retest reliability, equivalence, and internal consistency), and some limited evidence about generalizability across contexts (e.g.

psychological research, health care intake forms) in Chapter 4. In this chapter, I consider the generalizability of the multiple-choice gender identity question across participants.

John and Benet-Martinez (2000) suggest that there is no necessary or sufficient level of generalizability which a measure must achieve to be valid; rather, there are valid and invalid degrees of generalization from a particular measure. The designer of a given measure must therefore carefully consider the scope of the construct being studied, as well as the measure's intended use, when determining what level of generalizability is appropriate. The considerations involved in making this decision are often referred to as the bandwidth-fidelity tradeoff, which states that measures of broad constructs can predict a wide range of outcomes with low fidelity, while measures of narrow constructs can predict a narrow range of outcomes with much higher fidelity. For example, political ideology (liberal vs. conservative) is a broad construct because it predicts many outcomes with a low level of precision (there is substantial variation within each category). In contrast, a measure of attitudes toward tax cuts is a narrow construct which predicts specific behaviors related to tax cuts (e.g. voting decisions on a referendum) with much greater precision, but may not accurately predict anything about attitudes toward immigration.

In the context of gender questions, the ubiquitous binary gender question clearly measures (albeit imperfectly) an extremely broad construct, and can thus predict a wide range of outcomes. However, I argue that the fidelity of this question is low; there is a large amount of variation within each group. In contrast, the multiple-choice gender identity measure is designed to capture a relatively narrow construct with higher fidelity, and we feel that it is only appropriate to predict a limited range of outcomes (i.e. those related specifically to *identity*) and less appropriate to predict gender presentation, gender roles, etc.

That said, I do have a few specific predictions about the intended generalizability of the proposed measure which will need to be tested. The measure should be appropriate for use with both cisgender and transgender participants, meaning that conclusions drawn from cisgender participants' responses to the proposed measure should also be applicable to transgender people, and vice versa (see the sections on substantive validity and external validity for examples of such outcomes). The measure was also designed to capture both binary and non-binary gender identities, so both types of gender identities should have similar implications for other outcomes related to identity. To assess these, I included measures of gender centrality, gender identification, and emotions toward one's own gender group, which I expect to be similar within gender identities regardless of trans status.

### **External Validity**

External validity can be demonstrated by showing that a measure relates to other previously-established measures in the ways that would be predicted by theory. Thus, a measure with high external validity should be highly related to theoretically similar measures (e.g. other measures of the same construct) and mostly unrelated to theoretically dissimilar measures. In addition, an externally valid measure should be able to predict conceptually-related behaviors and outcomes.

I expect to find significantly different patterns of responding by gender identity (and in certain cases by transgender identity) to the Multi-Gender Identity Questionnaire (Joel et al., 2014) and my self-created Non-Binary Experiences Scale. For example, I expect both men and women to be more likely to feel like, use the pronouns of, and wear clothing consistent with their genders. Similarly, I expect both non-binary and agender people to be significantly less likely to

feel like men or like women, more likely to feel like neither a man nor a woman, more likely to use gender-neutral pronouns, and more likely to endorse non-binary experiences.

## **Overview**

I present evidence from two studies about the substantive validity, generalizability, and external validity of the multiple-choice gender identity measure, alongside the transgender identity measure. Substantive validity evidence comes from self-definition, self-investment, and experiences of identity denial. Generalizability evidence comes from gender centrality, gender identification, and emotions toward one's own gender group, and external validity evidence comes from pronoun usage, clothing choices, feeling like a man or woman, and endorsement of non-binary experiences.

## **Method**

This chapter and Chapters 3 and 4 describe results from different parts of the same two studies (each sampling undergraduates and self-identified gender minorities for four total samples). For a detailed description of the sample, the gender demographic measures, and the coding techniques used for each one, see Chapter 3.

After answering the gender demographic questions and indicating their perceptions of those questions and the researchers who asked them (see Chapters 3 and 4), participants completed a series of measures designed to provide preliminary evidence of the validity of the multiple-choice gender identity question (when paired with the transgender identity question).

## **Materials**

### **Gender Identity Measures.**

In both studies, participants completed multiple-choice gender identity questions and a transgender identity measure (Figures 1 and 2). The following analyses use the condensed

multiple-choice gender identities and transgender identities (see Table 1 for frequencies). For details about how these measures were constructed, see Chapter 2; for details about how they were coded, see Chapter 3.

Table 1

*Frequencies of condensed multiple-choice gender identity and condensed transgender identity in Study 1 and Study 2*

Condensed Transgender Identity	Condensed Multiple-Choice Gender Identity					
	Men		Women		Non-Binary/Agender	
	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2
Not Transgender	99	221	93	263	1	11
Transgender	38	12	42	12	133	35
Unclassified	7	9	7	6	4	7

**What is your gender identity?**

- Man
- Woman
- Genderqueer
- Another identity not listed (please specify)
- Do not know
- Choose not to answer

“Transgender/gender non-conforming” describes people whose gender identity or expression is different, at least part of the time, from the sex assigned to them at birth. **Do you consider yourself to be transgender/gender non-conforming in any way?**

- Yes
- No
- Do not know
- Choose not to answer

*Figure 1. Multiple-choice gender identity question and transgender identity question used in Study 1.*

What is your gender identity?

Woman

Man

None; I am agender.

Non-binary (please specify the term you use)

Another identity not listed (please specify the term you use)

Do not know

Choose not to answer

---

"Transgender" describes people whose gender identity or expression is different, at least part of the time, from the sex assigned to them at birth. **Do you consider yourself to be transgender?**

No

Yes

Do not know

Choose not to answer

Figure 2. Multiple-choice gender identity question and transgender identity question used in Study 2.

## Dependent Variables.

### *Substantive Validity – Self-Definition and Self-Investment.*

*Multi-Component Social Identity Scale (Leach et al., 2008; Study 1).<sup>19</sup>*

We adapted the Multi-Component Social Identity Scale (Leach et al., 2008) to ask about identity as a man and identity as a woman. These questions are as follows:

- *I feel a bond with [men/women].*
- *I feel solidarity with [men/women].*
- *I feel committed to [men/women].*
- *I am glad to be a [man/woman].*
- *I think that [men/women] have a lot to be proud of.*
- *It is pleasant to be a [man/woman].*
- *Being a [man/woman] gives me a good feeling.*
- *I often think about the fact that I am a [man/woman].*
- *The fact that I am a [man/woman] is an important part of my identity.*
- *Being a [man/woman] is an important part of how I see myself.*
- *I have a lot in common with the average [man/woman].*

<sup>19</sup> Participants were also asked to provide comments about the group identification questions.

- *I am similar to the average [man/woman].*
- *[Men/women] have a lot in common with each other.*
- *[Men/women] are very similar to each other.*

All participants answered all questions about identity as a man and all questions about identity as a woman in order to test the ability of each gender demographic question to predict identity as a man, woman, both, or neither (in line with Tate, Youssef, & Bettergarcia, 2014). However, the original Multi-Component Social Identity Scale is designed to be used only with participants who have already self-categorized into a particular group, so the wording of certain questions might have been confusing to participants. For example, the item “*Being a man is an important part of how I see myself*” might be confusing to women. In an attempt to address this, we provided participants with a “Not applicable” option in addition to the original 7-point Likert scales (1 = Strongly disagree, 7 = Strongly agree) used by Leach et al. (2008). All responses of “Not applicable” were omitted from these calculations. Both scales had high reliability (identity as man  $\alpha = .93$ ; identity as woman  $\alpha = .92$ ) and the items were averaged into composite scores (identity as man  $M = 3.8$ ,  $SD = 1.6$ ; identity as woman  $M = 4.5$ ,  $SD = 1.4$ ).

*Single-Item Social Identification Scales (Postmes et al., 2013; Study 1 Gender Minority Sample).*

We assessed identification with each of several different gender groups using versions of a single-item social identification scale designed by Postmes et al. (2013). Agreement with the statement “I identify with [my group]” captures most of the variance in the self-investment dimension of the Multi-Component Social Identity scale developed by Leach et al. (2008), as well as some of the variance in the self-definition dimension, making this single item a concise way to assess general identification with a group. Participants indicated their agreement with the



following seven statements about their identification with various gender groups on 7-point Likert scales (1 = strongly disagree, 7 = strongly agree).

- *I identify with gender minorities* (M = 5.9 , SD = 1.4)
- *I identify with transgender people* (M = 5.4, SD = 1.7)
- *I identify with gender non-conforming people* (M = 5.9, SD = 1.5)
- *I identify with people with non-binary genders* (M = 5.8, SD = 1.7)
- *I identify with men* (M = 3.67, SD = 2.1)
- *I identify with women* (M = 4.46, SD = 1.8)
- *I identify with cisgender people* (M = 2.74, SD = 1.8)

This measure was included in the survey that was distributed to gender minority participants as a pilot test, and was omitted from the survey completed by the student sample.

*Identity as a Man/Woman/Non-Binary Person (adapted from the Multi-Gender Identity Questionnaire, Joel et al., 2014; Study 2).*

These questions were adapted from the Multi-Gender Identity Questionnaire (Joel et al., 2014) to be more inclusive of non-binary gender identities. All participants answered all questions regardless of their gender identity or transgender status. Questions were prefaced with “In the past 12 months, have you...” and answered on a 5-point Likert scale (1 = Always, 5 = Never). We first asked about identity as a man and identity as a woman (8 questions):

- *...thought of yourself as a [man/woman]?*
- *...felt more like a [man than like a woman/woman than like a man]?*
- *...felt that you did not have a lot in common with [men/women]?*
- *...felt that it is/it would be better for you to live as a [man than as a woman/woman than as a man]?*

We then asked about identity as a non-binary person (4 questions):

- *...felt somewhere in between a woman and a man?*
- *...felt that you are neither a man nor a woman?*
- *...felt that you have nothing in common with men and nothing in common with women?*
- *...felt at times more like a man and at times more like a woman?*

We aggregated the four items about identity as a man ( $M = 2.8$ ,  $SD = 1.2$ ,  $\alpha = 0.7$ ), woman ( $M = 2.9$ ,  $SD = 1.2$ ,  $\alpha = 0.75$ ), and non-binary person ( $M = 1.5$ ,  $SD = 0.9$ ,  $\alpha = 0.84$ ), respectively, to create a composite score for each identity.

### ***Generalizability***

*Gender Centrality (Leach et al., 2008; Study 1 Gender Minority Sample).*

A general measure of gender centrality was adapted from the centrality subscale of the Multi-Component Social Identity Scale (Leach et al., 2008) to avoid specifying a particular gender. It included three items measured on a 7-point scale (1 = strongly disagree, 7 = strongly agree).

- *I often think about my gender.*
- *My gender is an important part of my identity.*
- *My gender is an important part of how I see myself.*

These three questions had acceptable reliability ( $\alpha = .84$ ), so they were averaged into a single composite score ( $M = 5.8$ ,  $SD = 1.3$ ). This measure was included in the survey that was distributed to gender minority participants as a pilot test, and was omitted from the survey completed by the student sample.

*Gender Identification (Egan & Perry, 2001; Study 2).*

Egan and Perry (2001) designed a multidimensional gender measure explicitly for use in interviews with children. Their scale had five dimensions: knowledge of which gender category one belongs to, feelings of typicality for that gender and contentedness with that gender assignment, felt pressure to behave according to gender stereotypes, and bias toward one's own gender group. While their original measure was designed for use only with participants who had already been classified as girls or as boys, we modified their measure to be answerable by a person of any gender. We further adapted the format to be comparable to the other measures

used in our survey. Half of participants saw the positive wordings of items 1, 2, and 5 and the negative wordings of items 3, 4, and 6 (version 1), while the other half of participants saw the inverse (version 2). Participants were asked to indicate their level of agreement with each of the following statements on a 7-point Likert scale (1= Strongly disagree, 7= Strongly agree).

*When you see the words “my gender,” think about your gender identity, regardless of the way you are perceived by others.*

- 1. I feel that my personality is similar to the personalities of most members of my gender/ I don't feel that my personality is similar to the personalities of most members of my gender.*
- 2. I feel I fit in with other members of my gender/ I don't feel that I fit in with other members of my gender.*
- 3. I feel I'm just like all other members of my gender/ I don't feel I'm just like all other members of my gender.*
- 4. I feel that the things I like to do in my spare time are similar to what most members of my gender like to do in their spare time/ I don't feel that the things I like to do in my spare time are similar to what most members of my gender like to do in their spare time.*
- 5. I feel that the kinds of things I'm good at are similar to what most members of my gender are good at / I don't feel that the kinds of things I'm good at are similar to what most members of my gender are good at.*
- 6. I think I am a good example of my gender/ I don't think I am a good example of my gender.*

All negatively worded items (beginning with “I don't”) were reverse-coded and the scores were aggregated into a single composite variable. Scores on each version were very similar (version 1:  $M = 4.9$ ,  $SD = 1.2$ ; version 2:  $M = 5.2$ ,  $SD = 1.1$ ), and reliabilities were high for both version 1 ( $\alpha = .83$ ) and version 2 ( $\alpha = .81$ ).

#### *Emotions toward Gender Groups (Studies 1 and 2).*

We asked participants to indicate the degree to which they feel each of four positive (Pride, Satisfaction, Compassion, and Empathy) and six negative (Anger, Irritation, Anxiety, Fear, Guilt, and Shame) emotions toward each of seven gender groups (in Study 1) or toward their own gender group (in Study 1 and 2). For each group, the items for positive emotions and negative emotions were averaged into composite scores (all  $\alpha > .87$  for positive emotions, all  $\alpha$

>.92; see Table C1). Due to space concerns, we only analyzed the emotions toward one's own gender group here.

### ***Substantive Validity.***

*Identity Denial* (adapted from the Multi-Gender Identity Questionnaire (Joel et al., 2014; Study 2).

These questions were adapted from the Multi-Gender Identity Questionnaire (Joel et al., 2014) to be more inclusive of non-binary gender identities. All participants answered all questions regardless of their gender identity or transgender status. Unless otherwise specified, questions were prefaced with “In the past 12 months, have you...” and answered on a 5-point Likert scale (1 = Always, 5 = Never). These items ask participants how they feel about the pronouns others use for them (3 questions) and about reporting a gender on official forms (2 questions). In addition to the 5-point Likert scale, we included a “Not relevant” option to accommodate participants who have never had the experience of being referred to by those pronouns / been required to report a particular gender on official forms. Participants who selected “Not relevant” were excluded from analyses on these questions.

- ...been upset when other people used female pronouns (she/her) to refer to you?
- ...been upset when other people used male pronouns (he/him) to refer to you?
- ...been upset when other people used gender-neutral pronouns (e.g. they/them, ze/hir, etc.) to refer to you?
- ...been bothered by having to check the box [‘F’ for female/ ‘M’ for male] on official forms (e.g. driver’s license, passport)?
- ...been bothered by the fact that any gender information is requested on official forms?

### ***External Validity.***

*Pronoun Use* (adapted from the Multi-Gender Questionnaire, Joel et al., 2014; Study 2).

We asked participants about which pronouns they use to refer to themselves. These questions were adapted from the Multi-Gender Identity Questionnaire (Joel et al., 2014) to be

more inclusive of non-binary gender identities. All participants answered all questions regardless of their gender identity or transgender status. Questions were prefaced with “In the past 12 months, have you...” and answered on a 5-point Likert scale (1 = Always, 5 = Never).

- ...used male pronouns (he/him) to refer to yourself?
- ...used female pronouns (she/her) to refer to yourself?
- ...used gender-neutral pronouns (e.g. they/them, ze/hir, etc.) to refer to yourself?

*Clothing Choices* (adapted from the *Multi-Gender Questionnaire*, Joel et al., 2014; Study 2).

We asked participants about what types of clothing they typically purchase and wear. All participants answered all questions regardless of their gender identity or transgender status. Questions were prefaced with “In the past 12 months, have you...” and answered on a 5-point Likert scale (1 = Always, 5 = Never).

- ...when you went into a department store to buy yourself clothing, did you shop mostly in a department labeled for [men/women]?
- ...have you worn the clothes typically associated with [men/women]?

We combined the two items about men’s clothing ( $M = 2.9$ ,  $SD = 1.6$ ,  $\alpha = 0.95$ ) and the two items about women’s clothing ( $M = 3.0$ ,  $SD = 1.7$ ,  $\alpha = 0.97$ ) into composite scores to simplify analysis.

### ***Additional Measures Not Analyzed.***

*Multi-Gender Identity Questionnaire* (adapted from Joel et al., 2014; Study 2).

In addition to the measures already described, our adaptation of the Multi-Gender Identity Questionnaire included 10 questions about participants’ feelings about being a man and being a woman (5-point Likert scales, 1= Always, 5 = Never, and a Not Relevant option), four questions about their experiences of physical dysphoria (5-point Likert scales, 1= Always, 5 = Never, and a

Not Relevant option), and two questions about whether they would like to be reborn as a man or a woman (Options were No, To some extent, To a large extent, Not Sure, and Don't Care).

#### Feelings About Being a Man/Woman

- *In the past 12 months, have you felt pressured by others to be a 'proper' [man/woman]?*
- *In the past 12 months, have you felt that you were not a 'real' [man/woman]?*
- *In the past 12 months, have you felt satisfied being a [man/woman]?*
- *In the past 12 months, have you felt that you have to work at being a [man/woman]?*
- *In the past 12 months, have you had the wish or desire to be a [man/woman]?*

#### Experiences of Physical Dysphoria

- *In the past 12 months, have you disliked your body because it is [male/female]?*
- *In the past 12 months, have you disliked your body because of its primary sex characteristics (e.g. genitals)?*
- *In the past 12 months, have you disliked your body because of its secondary sex characteristics (e.g. chest shape, hair growth patterns)?*

#### Desire to Be Reborn as a Man/Woman

- *If you could be reborn, would you like to be born as a [man/woman]?*

These 16 questions were included because they were part of the original scale, but as they were not directly relevant to our goal of validating a gender identity question, we did not analyze them further.

#### *Gender Essentialism (Study 2)*

We included a newly created 28-item gender essentialism scale with the other dependent variables to collect pilot data for scale construction and modification (see Chapter 6).

#### *Transgender Congruence Scale (Kozee et al., 2012) and Gender Identity Reflection and Rumination Scale (Bauerband & Galupo, 2014)*

The Transgender Congruence Scale and Gender Identity Reflection and Rumination Scale were included as a preliminary attempt to understand the experience of dysphoria among non-binary and agender participants, a question which is outside the scope of this dissertation.

Participants completed the 12-item Transgender Congruence Scale developed by Kozee et al. (2012) using a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree). This scale was designed for use with a transgender population and assesses the level of congruence felt between one's physical appearance and one's gender identity (e.g. *My outward appearance represents my gender identity*), as well as general comfort with one's gender identity (e.g. *I have accepted my gender identity*). Scores had high reliability ( $\alpha = .94$ ), and were aggregated into a single composite score ( $M = 5.8$ ,  $SD = 1.2$ ).

Participants completed the 15-item Gender Identity Reflection and Rumination Scale (Bauerband & Galupo, 2014) using a 4-point Likert scale to indicate how often they have particular thoughts about their gender identity (1 = Almost never, 4 = Almost Always). This scale measures the frequency of rumination on topics related to one's gender identity (e.g. *Think about things I cannot do because of my gender identity; Try to figure out what others think about my gender identity*). Scores had high reliability ( $\alpha = .95$ ), and were aggregated into a single composite score ( $M = 1.6$ ,  $SD = 0.6$ ).

## Results

For all comparisons, we used condensed versions of the multiple-choice gender identity (only men, women, and non-binary/agender participants; other responses were dropped for this analysis) and transgender identity questions (Transgender, Not Transgender, Don't Know/No Answer). For details about the construction of these variables, refer to Chapter 3.

### Evidence for Substantive Validity

#### **Evidence for Substantive Validity from Self-Definition and Self-Investment.**

##### *Multi-Component Social Identity Scale (Leach et al. 2008) – Study 1.*

The interaction of gender identity and transgender identity had a significant effect on identity as a man ( $F(4, 385) = 3.01, p = .018$ ) and identity as a woman ( $F(4, 391) = 3.07, p = .017$ ). There was also a main effect of gender identity on identity as a man ( $F(2, 392) = 60.04, p < .001$ ) and as a woman ( $F(2, 398) = 64.08, p < .001$ ), and a main effect of transgender identity on identity as a man ( $F(2, 437) = 66.65, p < .001$ ) and as a woman ( $F(2, 443) = 27.89, p < .001$ ). Means comparisons are presented in Table C1.

Overall, men identified more as men than did women or non-binary/agender people, women identified more as women than did men or non-binary/agender people, and non-binary/agender people identified less as men or women than did men or women.

Cisgender men identified the most as men, followed by cisgender women, transgender men, and men who did not know whether they identified as transgender. Transgender women, women who did not know whether they identified as transgender, and all non-binary people identified least as men. Cisgender and transgender women identified the most as women, followed by cisgender men and all those who did not know whether they identified as



transgender. Transgender and non-transgender non-binary/agender people identified least as women.

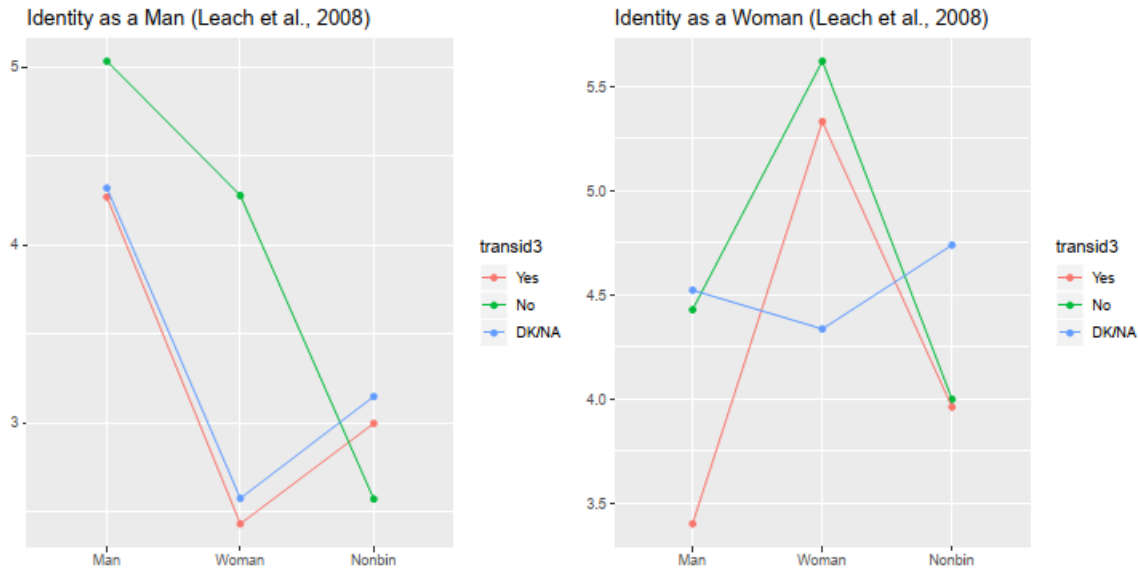


Figure 3. Mean identity as a man and mean identity as a woman (Leach et al., 2008) by gender identity and transgender identity (Study 1).

### *Single-Item Social Identification Scales (Postmes et al. 2013) – Study 1.*

These measures were only included in the gender minority survey. We could not analyze the interaction between gender identity and transgender identity for these variables because of an empty cell (there were no men in the sample who did not know whether they identified as transgender), so we report only the main effects.

#### *Identification with men.*

There was a main effect of gender identity ( $F(2, 222) = 55.66, p < .001$ ) on identification with men, but no effect of transgender identity. Men identified more as men than did women or non-binary/agender people.

#### *Identification with women.*

There was a main effect of gender identity ( $F(2, 224) = 48.17, p < .001$ ) on identification with women, but no effect of transgender identity. Women identified more as women than did men or non-binary/agender people.

*Identification with non-binary people.*

There was a main effect of gender identity ( $F(2, 223) = 54.62, p < .001$ ) and transgender identity ( $F(2, 267) = 28.28, p < .001$ ) on identification with non-binary people. Non-binary/agender people identified more as non-binary than did men or women, but transgender men and all women regardless of transgender identity identified with non-binary people next most strongly, while cisgender men had the lowest levels of identification with non-binary people.

*Identification with gender non-conforming people.*

There was a main effect of gender identity ( $F(2, 224) = 30.71, p < .001$ ) and transgender identity ( $F(2, 268) = 29.52, p < .001$ ) on identification with gender non-conforming people. Non-binary people, regardless of transgender identity, identified most strongly with gender non-conforming people. Transgender men and all women regardless of transgender identity identified next most strongly with gender non-conforming people. Cisgender men had the lowest levels of identification with gender non-conforming people.

*Identification with gender minorities.*

There was a main effect of gender identity ( $F(2, 224) = 8.70, p < .001$ ) and transgender identity ( $F(2, 268) = 37.59, p < .001$ ) on identification with gender minorities. All non-binary/agender people and all transgender people identified most strongly with gender minorities, followed by cisgender women and women who did not know whether they identified as transgender. Cisgender men had the lowest levels of identification with gender minorities.

*Identification with transgender people.*

There was no effect of gender identity on identification with transgender people, but there was a main effect of transgender identity ( $F(2, 267) = 27.92, p < .001$ ). All transgender people and all non-binary/agender people (regardless of transgender identity) identified more strongly with transgender people than cisgender men, cisgender women, or women who did not know whether they identified as transgender.

*Identification with cisgender people.*

There was a main effect of gender identity ( $F(2, 223) = 6.61, p = .002$ ) and transgender identity ( $F(2, 267) = 15.72, p < .001$ ) on identification with cisgender people. Cisgender men and cisgender women identified most strongly with cisgender people. All other categories identified less strongly with cisgender people.

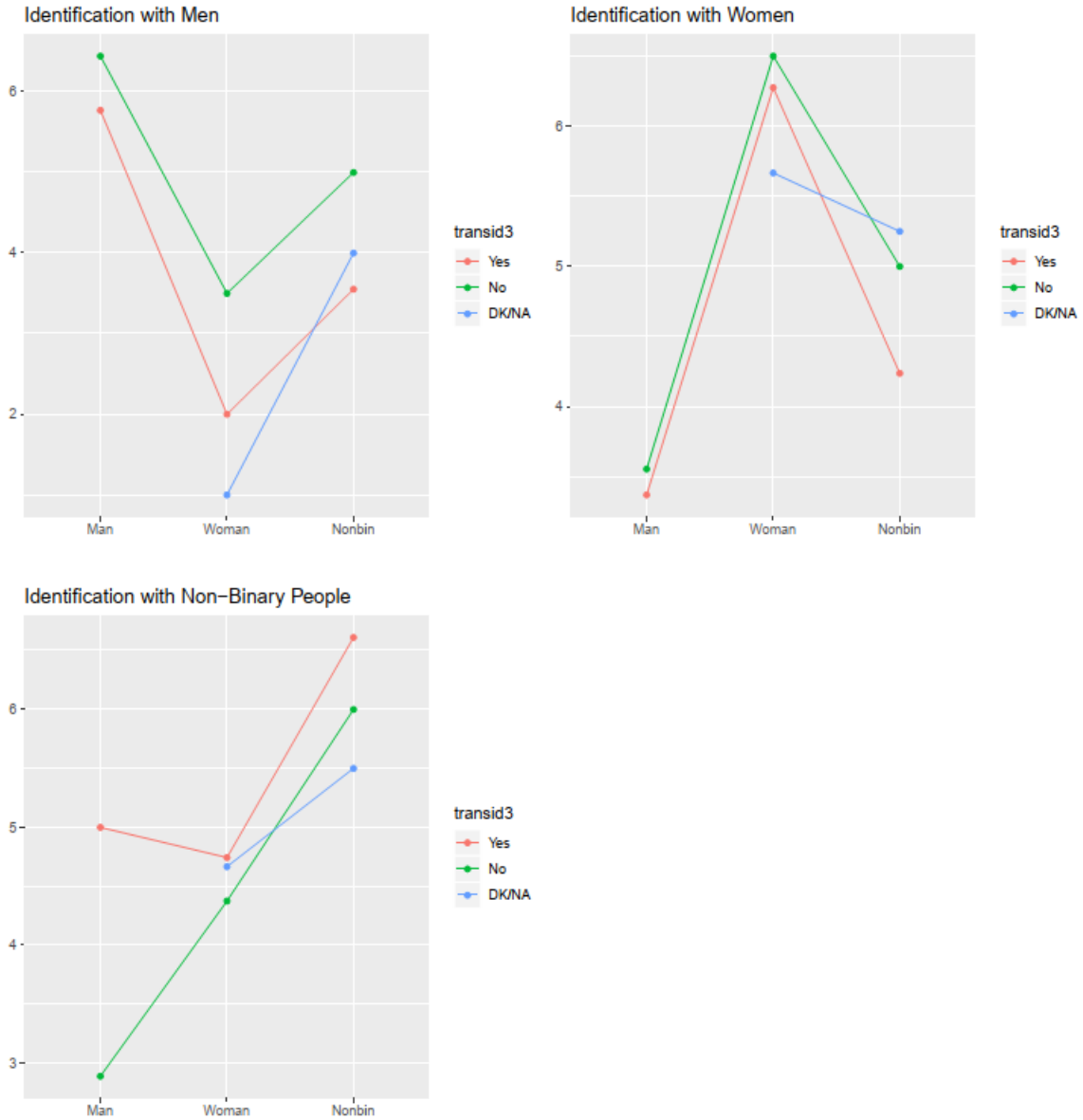


Figure 4. Mean identification with men, women, and non-binary people (Postmes et al., 2013) by gender identity and transgender identity (Study 1).

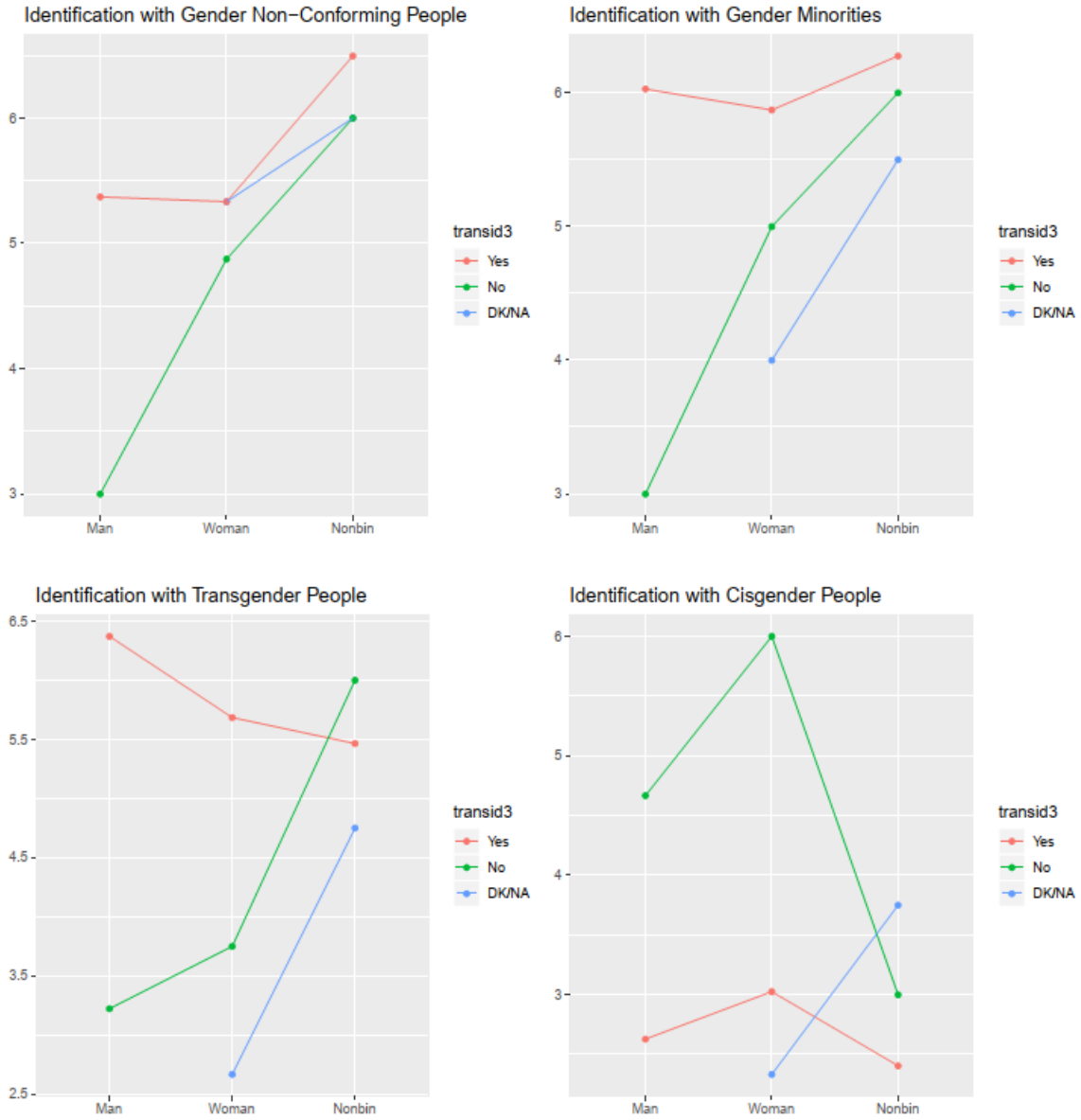


Figure 5. Mean identification with gender non-conforming people, gender minorities, transgender people, and cisgender people (Postmes et al., 2013) by gender identity and transgender identity (Study 1).

### ***Identity as Man/Woman/Non-Binary – Study 2.***

The interaction of gender identity and transgender identity had a significant effect on identity as a man ( $F(4, 544) = 3.47, p = .008$ ) and as a woman ( $F(4, 540) = 2.86, p = .023$ ), but not on identity as a non-binary person. Men who were unclassified on the transgender identity measure identified less as men than transgender or non-transgender men, and unclassified women identified less as women than transgender or non-transgender women. Transgender and non-transgender men did not differ in identification as men, and transgender and non-transgender women did not differ in identification as women.

There was a main effect of gender identity on identity as a man ( $F(2, 554) = 582.04, p < .001$ ), woman ( $F(2, 550) = 533.34, p < .001$ ), and non-binary person ( $F(2, 550) = 218.37, p < .001$ ). Men identified more as men than did women or non-binary/agender people. Women identified more as women than did men or non-binary/agender people. Non-binary/agender people identified more as non-binary than did women or men.

There was no main effect of transgender identity on identity as a man or as a woman, but there was a main effect of transgender identity on identity as a non-binary person ( $F(2, 559) = 118.89, p < .001$ ). Transgender and unclassified people identified more as non-binary people than non-transgender people.

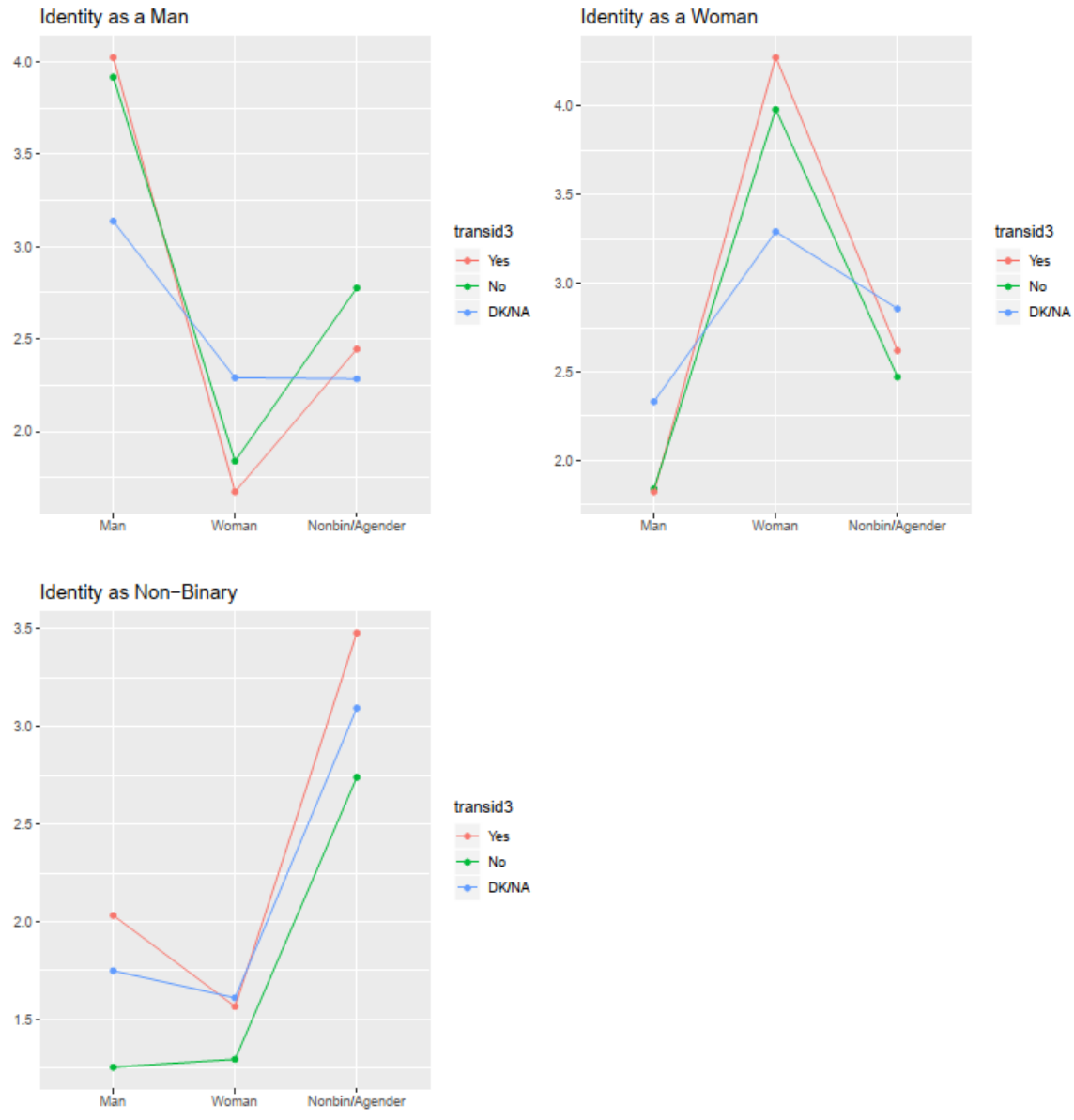


Figure 6. Mean identity as a man, a woman, or a non-binary person by gender identity and transgender identity (Study 2).

## **Evidence of Substantive Validity from Identity Denial.**

### **Discomfort with Mispronouncing**

The interaction of gender identity and transgender identity had a significant effect on being upset when others use he/him pronouns ( $F(4, 305) = 2.98, p = .019$ ), and she/her pronouns for you ( $F(4, 374) = 5.08, p = .001$ ). However, the interaction did not have a significant effect on being upset when others use gender neutral pronouns (they/them, ze/hir) for you.

There was a main effect of gender identity on being upset when others use he/him pronouns ( $F(2, 314) = 24.81, p < .001$ ), she/her pronouns ( $F(2, 383) = 107.89, p = .001$ ), and gender neutral pronouns for you ( $F(2, 319) = 5.82, p = .003$ ). There was a main effect of transgender identity on being upset when others use he/him pronouns ( $F(2, 319) = 12.48, p < .001$ ) and she/her pronouns ( $F(2, 392) = 55.33, p < .001$ ) for you. However, there was no significant effect of transgender identity on being upset when others use gender neutral pronouns for you. Means comparisons are presented in Tables C44, C48, and C52.

Transgender women and non-binary/agender people who did not know whether they identified as transgender were much more upset than any other groups when others used male pronouns to refer to them. All women and non-binary/agender people were more upset than men when others used male pronouns to refer to them, and transgender people were more upset than cisgender people.

Transgender men and transgender non-binary/agender people were most upset when others used female pronouns to refer to them. All men and non-binary/agender people were more upset than women, and transgender people were more upset than cisgender people.

Men (especially transgender men) were more upset than women or non-binary/agender people when others used gender neutral pronouns for them.



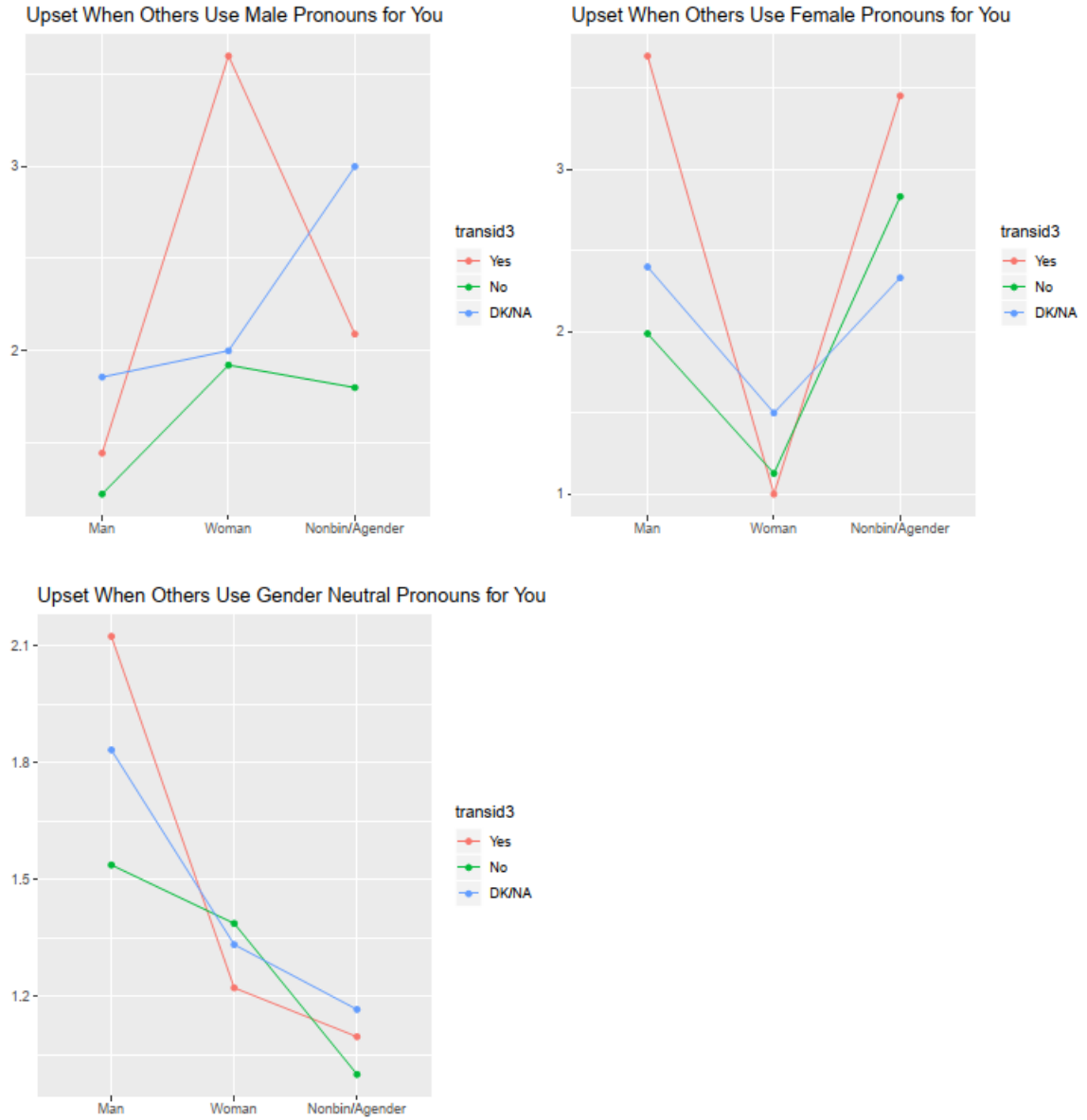


Figure 7. Mean frequency of being upset when others use male, female, or gender neutral pronouns for you by gender identity and transgender identity (Study 2).

### ***Discomfort with Gender Collection on Forms – Study 2.***

The interaction of gender identity and transgender identity had a significant effect on discomfort with checking the box “F” on official forms ( $F(4, 359) = 8.30, p < .001$ ), and on discomfort when any gender information is requested on official forms ( $F(4, 487) = 5.54, p < .001$ ). Because of an empty cell, the interaction could not be analyzed for the effect of discomfort with checking the box “M” on official forms. There was a main effect of gender identity on discomfort with checking the box “F” ( $F(2, 369) = 152.90, p < .001$ ), with checking the box “M” ( $F(2, 307) = 17.35, p < .001$ ), and with any gender information being requested on official forms ( $F(2, 497) = 115.77, p < .001$ ). There was also a main effect of transgender identity on discomfort with checking the box “F” ( $F(2, 375) = 100.59, p < .001$ ), with checking the box “M” ( $F(2, 310) = 34.26, p < .001$ ), and with any gender information being requested on official forms ( $F(2, 505) = 130.41, p < .001$ ). Means comparisons are presented in Tables C55, C59, and C63.

All transgender people (regardless of gender identity) had great discomfort that any gender information is collected on official forms, as did all non-binary/agender people (regardless of transgender identity). Non-binary/agender people (regardless of transgender identity) and transgender men experienced much greater discomfort with checking the box “F” on official forms than women (regardless of trans identity) or cisgender men. Transgender women and non-transgender non-binary/agender people had the most discomfort with checking the box “M” on official forms, but transgender non-binary/agender people and transgender men also had greater discomfort than cisgender men and women or men and women who did not know.

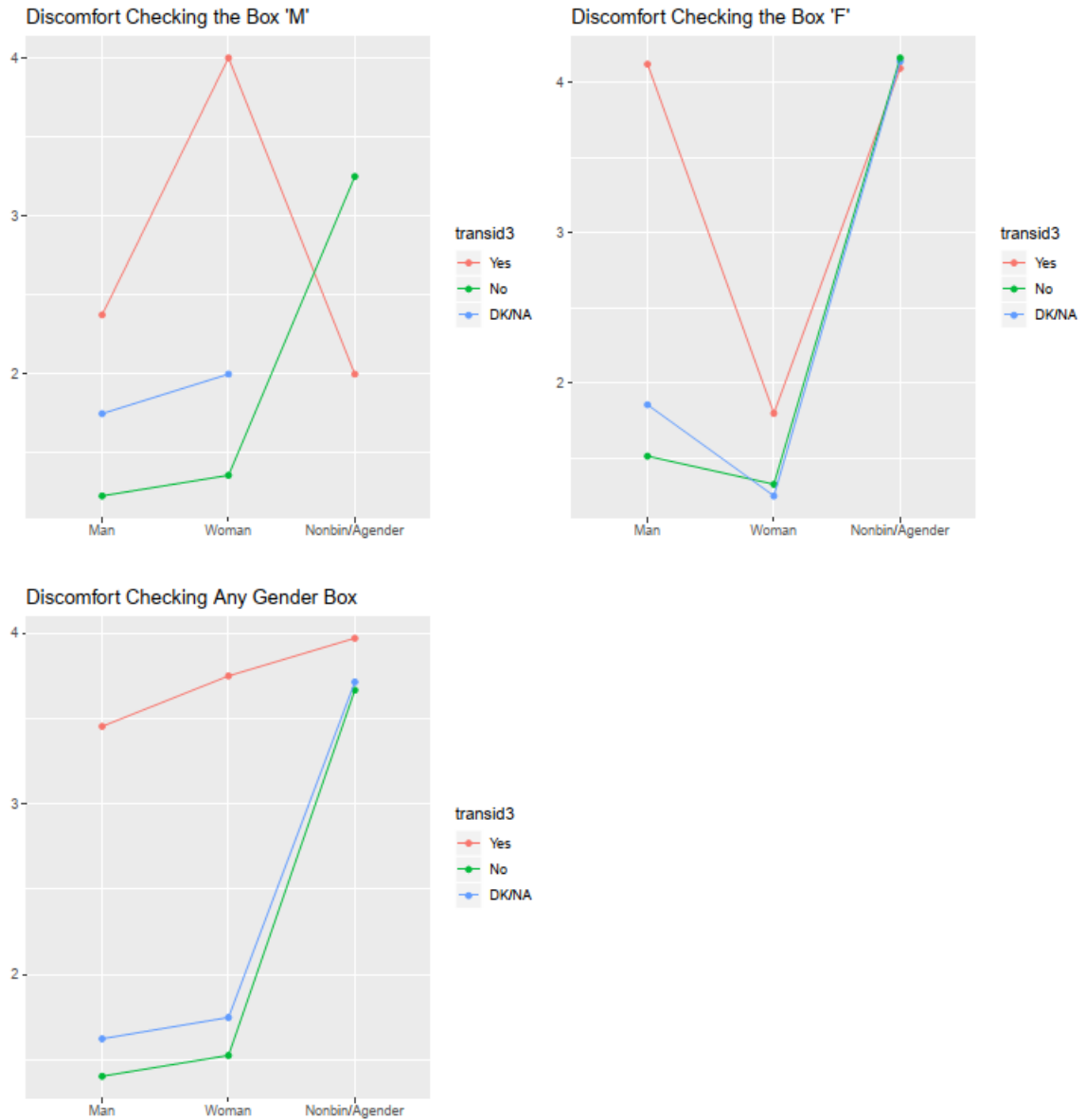


Figure 8. Mean frequency of discomfort when required to check the “M” or “F” box on official forms, or when any gender information is requested, by gender identity and transgender identity (Study 2).

## **Evidence for Generalizability.**

### ***Emotions Toward Own Gender Group – Study 1.***

The interaction of gender identity and transgender identity did not have a significant effect on positive or negative emotions toward one's own gender group. However, there was a main effect of gender identity on positive ( $F(2, 408) = 54.13, p < .001$ ) and negative emotions ( $F(2, 407) = 4.86, p = .008$ ), and a main effect of transgender identity on positive ( $F(2, 453) = 31.56, p < .001$ ) and negative emotions ( $F(2, 452) = 4.17, p = .016$ ).

Women had more positive emotions toward their own gender group than did men or non-binary people, and transgender people had more positive emotions toward their own gender group than did cisgender or unclassified people, regardless of gender identity. People who were unclassified had more negative emotions toward their own gender group than did cisgender or transgender people.

### ***Emotions Toward Own Gender Group – Study 2.***

The interaction of gender identity and transgender identity had a significant effect on positive emotions toward one's own gender group ( $F(4, 543) = 2.52, p = .040$ ), but not on negative emotions. There was a main effect of gender identity on positive ( $F(2, 553) = 47.05, p < .001$ ) and negative emotions ( $F(2, 547) = 5.40, p = .005$ ) toward one's own gender group. There was also a main effect of transgender identity on positive ( $F(2, 562) = 21.59, p < .001$ ) and negative emotions ( $F(2, 556) = 3.40, p = .034$ ) toward one's own gender group.

Transgender people had higher levels of positive emotion toward their own gender groups than cisgender people or those who did not know, and non-binary/agender people had higher levels of positive emotion toward their own gender groups than men or women. However,

cisgender men had much lower positive emotions toward their gender group than transgender men or those who did not know.

Nonbinary/agender people had the lowest levels of negative emotion toward their gender group, regardless of transgender status. Unclassified men and women had the highest levels of negative emotions toward their own gender group, and cisgender and transgender men and women did not differ.

### ***Gender Centrality – Study 1.***

The gender centrality questions were only included in the gender minority survey, so we could not analyze the effect of the interaction of gender identity and transgender identity due to empty cells in the analysis. There was no main effect of gender identity on gender centrality, but there was a main effect of transgender identity on gender centrality ( $F(2, 268) = 19.81, p < .001$ ) such that transgender participants were higher on gender centrality compared to cisgender and unclassified participants.

### ***Gender Identification – Study 2.***

The interaction of gender identity and transgender identity had a significant effect on gender identification ( $F(4, 548) = 4.63, p = .001$ ). There was also a main effect of gender identity ( $F(2, 558) = 20.57, p < .001$ ) and transgender identity ( $F(2, 567) = 35.04, p < .001$ ) on gender identification. Non-transgender men and women had much higher gender identification than non-transgender non-binary/agender people. Transgender and unclassified people did not differ in their level of gender identification by gender.

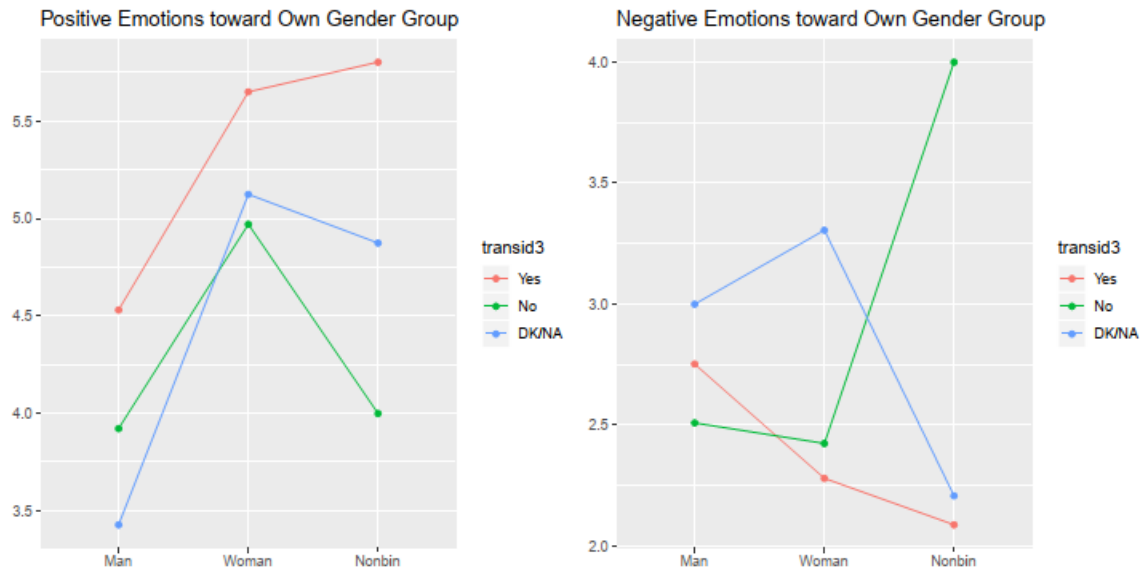


Figure 9. Mean positive and negative emotions toward one's own gender group by gender identity and transgender identity (Study 1).

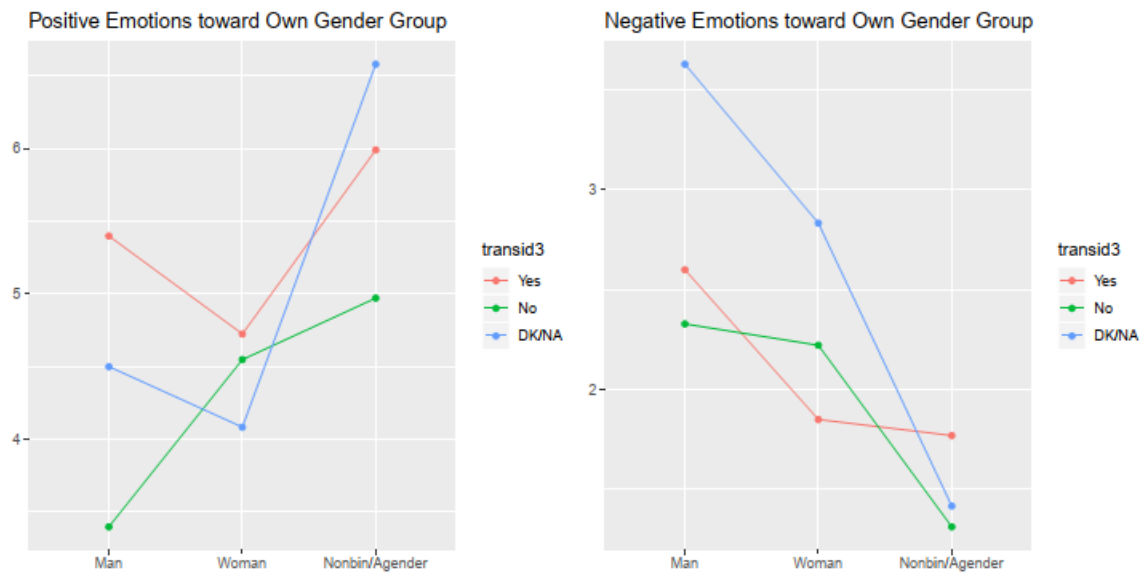


Figure 10. Mean positive and negative emotions toward one's own gender group by gender identity and transgender identity (Study 2).

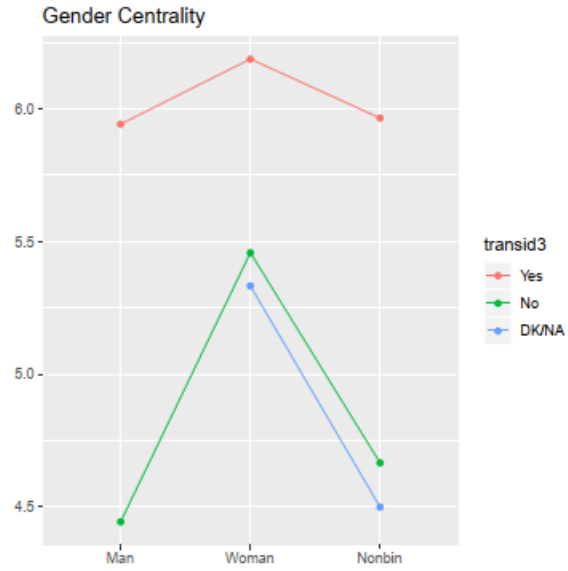


Figure 11. Mean gender centrality by gender identity and transgender identity (Study 1).

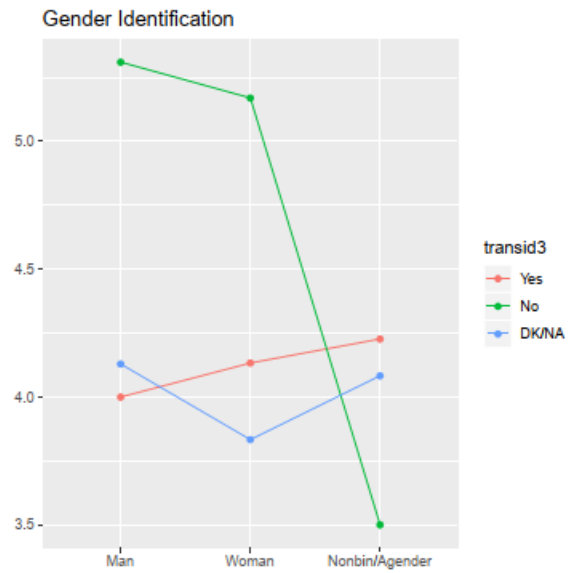


Figure 12. Mean gender identification by gender identity and transgender identity (Study 2).

## **Evidence for External Validity**

### *Use of Pronouns – Study 2.*

The interaction of gender identity and transgender identity could not be analyzed for pronoun use because of empty cells. There was a significant main effect of gender identity on use of male pronouns (he/him) ( $F(2, 559) = 455.46, p < .001$ ), female pronouns (she/her) ( $F(2, 559) = 536.67, p < .001$ ), and gender neutral pronouns (they/them, ze/hir) ( $F(2, 560) = 116.69, p < .001$ ) to refer to yourself. There was a no significant main effect of transgender identity on use of male pronouns (he/him) or female pronouns, but there was a significant effect on use of gender neutral pronouns (they/them, ze/hir) to refer to yourself ( $F(2, 569) = 61.83, p < .001$ ).

Men used male pronouns to refer to themselves more often than women or non-binary/agender people did. Transgender women used male pronouns for themselves more often than non-transgender women and those who did not know. Non-transgender non-binary/agender people used male pronouns for themselves more often than transgender non-binary/agender people or those who did not know.

Women used female pronouns to refer to themselves more often than men or non-binary/agender people did. Cisgender women used female pronouns for themselves more than transgender women or those who did not know. Men rarely used female pronouns for themselves, regardless of trans identity.

Non-binary/agender people used gender neutral pronouns to refer to themselves more often than men or women did. Transgender non-binary/agender people used gender neutral pronouns to refer to themselves more often than non-transgender non-binary/agender people, or those who did not know.



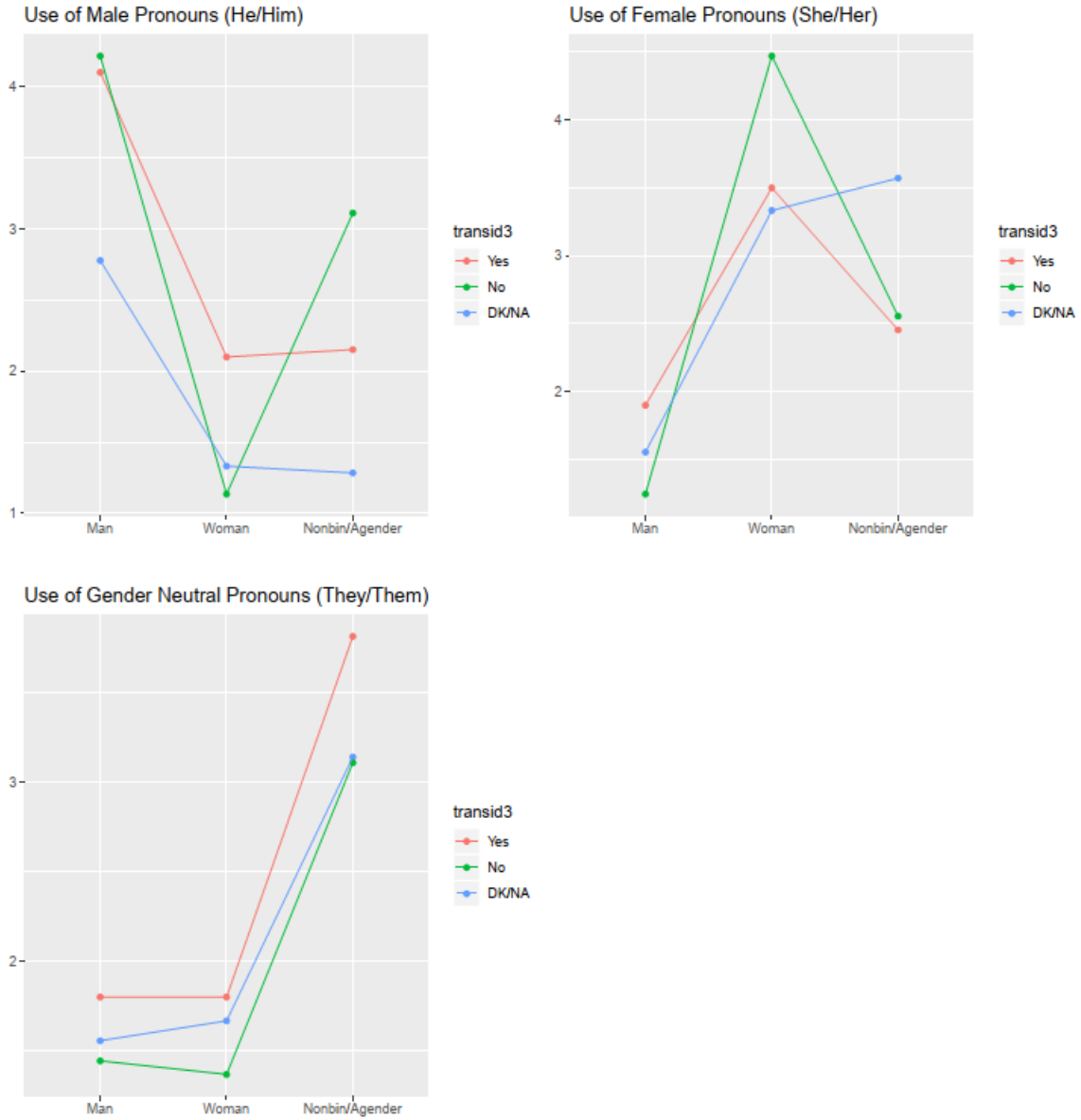


Figure 13. Mean frequency of using male, female, or gender neutral pronouns to refer to yourself by gender identity and transgender identity (Study 2).

### ***Clothing Choices – Study 2.***

The interaction of gender identity and transgender identity was not significant for choosing men's clothing, but was significant for choosing women's clothing ( $F(4, 551) = 9.57, p < .001$ ). There was no main effect of transgender identity on choosing men's or women's clothing. There was a main effect of gender identity on choosing men's clothing ( $F(2, 558) = 30.34, p < .001$ ) and women's clothing ( $F(2, 561) = 652.76, p < .001$ ).

Men were more likely to choose men's clothing than women were, and women were more likely to choose women's clothing than men were. Cisgender women were more likely to choose women's clothing than transgender women or those who did not know their transgender identity. Non-binary/agender people who did not know whether they identified as transgender were more likely to wear women's clothing than transgender or non-transgender non-binary/agender people. Men were unlikely to choose women's clothing regardless of transgender status.

### ***Non-Binary Experiences Scale – Study 2.***

The interaction of gender identity and transgender identity had no effect on endorsement of non-binary experiences. However, there was a main effect of gender identity ( $F(2, 304) = 37.81, p < .001$ ) and transgender identity ( $F(2, 308) = 36.58, p < .001$ ). Non-binary/agender people endorsed non-binary experiences more frequently than men or women. Those who were unclassified also endorsed non-binary experiences more frequently than cisgender or transgender people.

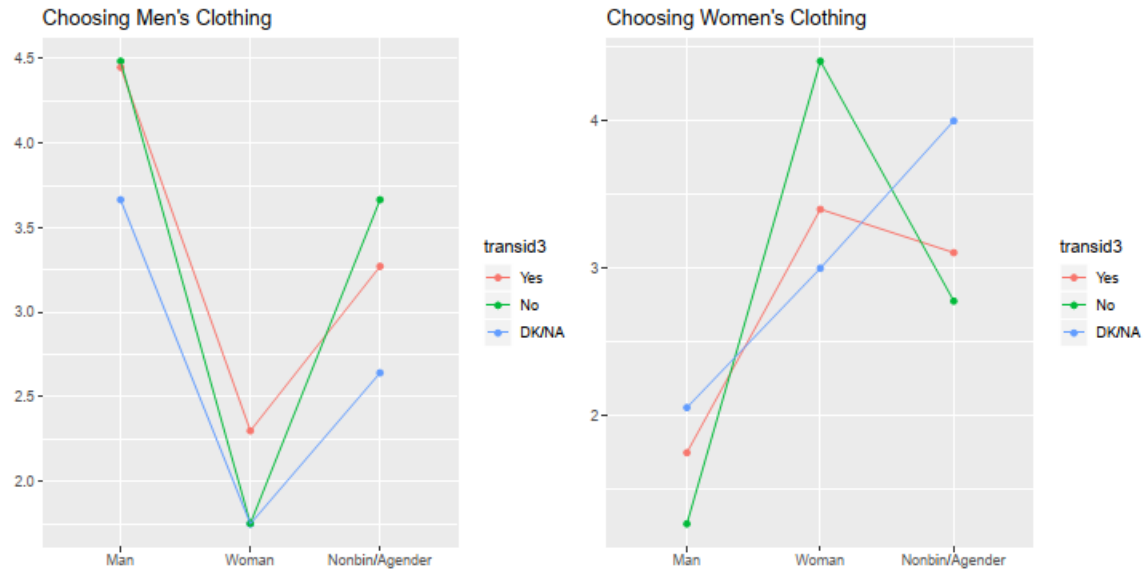


Figure 14. Mean frequency of choosing men's and women's clothing by gender identity and transgender identity (Study 2).



Figure 15. Mean scores on Non-Binary Experiences Scale by gender identity and transgender identity (Study 2).

## **Discussion**

As described in Chapters 3 and 4, the multiple-choice gender identity question combined with a question about transgender identity provided the best combination of theoretical precision, simplicity of analysis, and acceptability to participants. This chapter presented evidence for the substantive validity, generalizability, and external validity of the multiple-choice gender identity measure, heavily based on discussions of validation research by Messick et al. (1989) and John and Benet-Martinez (2000). Overall, the multiple-choice gender identity measure seems suitable for use with both cisgender and transgender populations.

### **Evidence of substantive validity**

To determine the substantive validity of the multiple-choice gender identity question, we tested its ability to predict levels of self-definition and self-investment in different gender groups, as well as the experience of identity denial from pronoun use or gender selection on official forms. We expected participants to experience greater self-investment and self-definition in the gender groups into which they categorized themselves, and to experience identity denial when being referred to by the wrong gender pronoun (“mispronouncing”) or being required to indicate a gender other than the one they identify with on an official form. These hypotheses were supported, suggesting that the multiple-choice gender identity question has good substantive validity for use in both cisgender and transgender populations, and with non-binary/agender people.

### **Evidence from self-definition and self-investment.**

#### **Multi-Dimensional Social Identity Scales (Leach et al., 2008), Study 1**

The Leach et al. (2008) Multi-Dimensional Social Identity Scales (one for identity as a man and one for identity as a woman) showed the expected patterns overall, with men

identifying more as men than did women or non-binary/agender people, women identifying more as women than did men or non-binary/agender people, and non-binary/agender people identifying less as men or women than did men or women. However, there were some unexpected differences by transgender identity. Transgender people showed the expected pattern of high identification with their own gender and lower identification with the other genders, but while cisgender men and women each identified more strongly with their own genders, the difference between cisgender men and cisgender women's identification with men was smaller than we expected. We attribute this to the unfamiliarity of being asked about level of identification with a gender category to which one does not belong; the Multi-Dimensional Social Identity Scale (Leach et al., 2008) explicitly assumes that a person completing it already self-identifies as a member of that group, so these results, while suggestive, should be interpreted with caution.

#### **Single-Item Social Identification Scales (Postmes et al. 2013), Study 1**

Responses to the Single-Item Social Identification Scales (Postmes et al. 2013) looked more as we expected. Men, women, and non-binary/agender people each identified more with their own gender than the other two genders did. Non-binary/agender people identified somewhat more as men than women did, and somewhat more as women than men did. These patterns were identical regardless of transgender identities. All non-binary/agender people (regardless of transgender identity) and all transgender people (regardless of gender identity) identified more with gender non-conforming people, gender minorities, and transgender people than did cisgender men, cisgender women, or women who did not know whether they identified as transgender. Conversely, cisgender men and cisgender women identified most strongly with cisgender people, and all other categories identified less strongly.

The fact that all non-binary/agender people (regardless of transgender identity) identified strongly with transgender people and very weakly with cisgender people is somewhat surprising. It suggests that the identity “non-binary” may connote an affiliation with transgender identities and a distancing from cisgender ones, rather than simply a gender identity (and/or that the transgender identity question is a very poor predictor of self-definition and self-investment in transgender groups for non-binary/agender people!). Future research should examine the relationship between non-binary/agender identities and the label “transgender” to shed further light on what non-binary and agender people mean when they do or do not choose to call themselves transgender.

### **Identity as a Man/Woman/Non-Binary Person (Joel et al., 2014), Study 2**

Men, women, and non-binary/agender people each identified more as their own gender than did the other two genders. Transgender and cisgender men and women did not differ in their identities as men or women, but men and women who were unsure whether they were transgender had lower identification with their gender group (as would be expected given that uncertainty about transgender status probably indicates uncertainty about gender identity membership, at least for men and women). Transgender and unclassified people, as well as non-binary/agender people regardless of transgender identity, identified more as non-binary (presumably because transgender status is often socially construed as “third-gender” membership).

These results suggest that responses to the multiple-choice gender identity measure do predict self-definition and self-investment in those gender groups. This was true regardless of transgender identity; that is, the multiple-choice gender identity question predicts level of identification with a gender category for both transgender and cisgender people of that gender,

and both transgender and cisgender people tend to identify equally strongly with their genders. Non-binary/agender people also had a distinct response pattern which was clearly distinguishable from those of men and women, suggesting that non-binary/agender identity can be meaningfully measured.

### **Evidence from identity denial.**

#### **Mispronouncing (Joel et al., 2014) Study 2**

As anticipated, women were more upset than men or non-binary people when others used male pronouns for them, and men were more upset than women or non-binary people when others used female pronouns for them. This was especially true for transgender people, who presumably both experience mispronouncing more frequently and feel greater identity denial when it occurs. Interestingly, while men were somewhat more upset about others using gender neutral pronouns for them than women or non-binary/agender people were, overall levels of discomfort were quite low for all participants. This may reflect the fact that use of the singular gender neutral pronoun “they” is increasingly common in English (Guo, 2016; Baron, 2018).

#### **Gender Collection on Official Forms (Joel et al. 2014) Study 2**

We also found that responses to the multiple-choice gender identity question predicted the experience of discomfort when selecting identity-denying gender options on official forms, but this pattern only held for transgender people (presumably because cisgender people do not generally have this experience). More generally, all transgender people and non-binary/agender people had great discomfort that any gender information is collected on official forms. In addition to the risk of identity denial, “official” connotes legal documents carrying a risk of financial, housing, employment, or health care discrimination against these gender minorities if their response is deemed unsuitable by the entity evaluating it.

Interestingly, while transgender women and non-transgender non-binary/agender people had the most frequent discomfort when checking the identity-denying box “M” on official forms, transgender non-binary/agender people and transgender men still had more discomfort than cisgender men and women or men and women who did not know whether they were transgender. The simplest explanation for this unexpected pattern is a combination of double-barreled survey wording and sample characteristics. In the original question “*In the past 12 months, have you been bothered by having to check the box [‘F’ for female/ ‘M’ for male] on official forms (e.g. driver’s license, passport)?*” we cannot tell with certainty whether a response indicates how often a person feels discomfort when an event occurs, or how often the event occurs in the first place. Because the sample of transgender non-binary/agender people disproportionately consisted of people assigned female at birth (N = 3 AMAB vs. N = 26 AFAB), while the non-transgender non-binary/agender people were roughly evenly split (N= 4 AMAB vs. N = 5 AFAB), the transgender non-binary/agender people were less likely to be placed in the position of needing to check the box “M” in the first place. This presumably caused the unexpected low frequency of transgender non-binary/agender people experiencing discomfort when checking the box “M” on official forms.

However, for transgender men, this explanation seems insufficient. Instead, it may be that transgender men do feel more uncomfortable indicating their membership in the category “men,” even though it is identity affirming. This interpretation is supported by a variety of qualitative comments from trans men across the two surveys. For example, Study 1 included space for free-text comments following the Leach et al. scales, and several participants left comments about an ambivalent relationship to the category man:

“Given the prevalence of misogyny and all kinds of sexism across our society and the ways that the idea of “man” is constructed in opposition to “woman” in part for the



purpose of subjugating women, it's difficult to have an uncomplicated relationship to the identity "man". That said, for all intents and purposes, I live as a man, and feel far more uncomfortable identifying as anything else, given that I am a recipient of male privilege. I do feel a certain bond to gay maleness in particular, and to a lesser extent to gender-nonconforming maleness."

"My idea of man is very traditional and hard for me to resonate with due to the toxic masculinity."

"I have desired to be in this male box my entire life, but I have always had to fight for it, I feel comfortable with these questions but also somewhat guilty since I am the only one that thinks this is the box I belong in, I feel like all the people born into this box don't want me there."

This suggests that transgender men may often have a less positive view of masculinity than cis men, whether due to personal experiences with misogyny, or a greater understanding of intersectionality of identities.

Overall, the evidence from identity denial partially supports our hypotheses. Responses to the multiple-choice gender identity question do predict discomfort when being mispronounced (for binary gender pronouns, but not for gender neutral ones). Likewise, they do predict discomfort when entering the wrong gender identity information on official forms (for transgender people).

## **Generalizability**

### **Gender Centrality and Identification**

Gender centrality was much higher for transgender people than for cisgender people, regardless of gender identity, but did not differ by gender identity. This measure was only administered to the gender minority sample in Study 1, so the sample size of non-transgender people was quite small. Given the extreme difficulties many transgender people face in relation to their gender identity, it is unsurprising that gender identity would then be quite central to them.

Conversely, gender identification (Egan & Perry, 2001) was much higher for cisgender men and cisgender women than any other group (the others did not differ). This is presumably because items from this scale measure feeling similar to conventional members of one's own group (rather than identification in the sense used by Leach et al. 2008). While we did not specifically anticipate this outcome, it makes sense that gender minorities would feel different from other members of their gender, as the prototypical member of those categories is cisgender. In contrast, gender minorities by definition must go through a process of questioning their birth assignment to a binary gender category, presumably creating a feeling of dislocation.

### **Emotions Toward Own Gender Group (Studies 1 and 2)**

In both Study 1 and Study 2, transgender people had higher levels of *positive emotion* toward their own gender groups than cisgender people or those who did not know whether they identified as transgender. In Study 1, women had higher levels of positive emotion toward their own gender groups than men, but did not differ from non-binary/agender people. In Study 2, non-binary/agender people had higher levels of positive emotion toward their own gender groups than men or women, and cisgender men had much lower positive emotions toward their gender group than transgender men or those who did not know whether they identified as transgender.

In both Studies 1 and 2, nonbinary/agender people had the lowest levels of *negative emotion* toward their gender group, regardless of transgender status. Unclassified men and women had the highest levels of negative emotions toward their own gender group, and cisgender and transgender men and women did not differ.

While it seems clear that transgender people and non-binary/agender people generally feel quite positively/not negatively toward their own gender groups, cisgender men generally had very low positive emotions and high negative emotions. Cisgender men may be less likely than

cisgender women or transgender people in general to include gender in their self-concepts because of their status as the default, privileged category, which may mean that being a man in and of itself does not produce the positivity that comes from membership in a marginalized/stigmatized group (e.g. Graham-Bailey et al., 2018). Growing cultural attention to sexism, toxic masculinity, and rape culture may also lead many men to feel discomfort or ambivalence about a male identity, as suggested by some of the comments presented earlier in this chapter. For example, our data collection throughout 2016 was bookended by the circa-2014 #notallmen/#yesallwomen furor (Zimmerman, 2014), and the more recent #metoo and #timesup movements (Langone, 2018).

### **External Validity**

To determine whether the multiple-choice gender identity measure has external validity, we tested its ability to predict behaviors and outcomes that are conceptually related to gender identity, but that are not constitutive of it. That is, we expected men and women to be more likely to feel like, use the pronouns of, and wear clothing consistent with their genders, and both non-binary and agender people to be significantly less likely to feel like men or like women, more likely to feel like neither a man nor a woman, and more likely to use gender-neutral pronouns. These hypotheses were supported.

#### **Pronoun Use (Joel et al. 2014, Study 2)**

Men used male pronouns, women used female pronouns, and non-binary/agender people used gender-neutral pronouns to refer to themselves more often than the other two genders did supporting the ability of the multiple-choice gender identity question to predict pronoun usage. There were some differences in pronoun usage by transgender identity; transgender women were more likely to use male pronouns and less likely to use female pronouns for themselves

compared to cisgender women, but there was no such difference by transgender identity among men. This suggests that transgender men may present as men in more aspects of their lives than transgender women do (perhaps because of the increased risk of discrimination and violence experienced by transgender women). Transgender non-binary/agender people were more likely than non-transgender non-binary/agender people to use gender-neutral pronouns for themselves. Possibly, non-binary/agender people who have taken specific behavioral steps to publicly express their identities (such as pronoun usage) may be more likely to identify as transgender.

### **Clothing Choices (Joel et al. 2014, Study 2)**

Men and women were each more likely to choose clothing associated with their gender than the other genders were, supporting the ability of the multiple-choice gender identity question to predict clothing choices. However, while men were unlikely to choose women's clothing regardless of transgender status, cisgender women were more likely to choose women's clothing than transgender women or those who did not know their transgender identity. Transgender non-binary/agender people were somewhat less likely to choose women's clothing than those with other transgender identities, but overall, non-binary/agender did not exhibit a clear pattern of gendered clothing choices.

### **Non-Binary Experiences Scale – Study 2.**

Non-binary/agender people endorsed non-binary experiences more frequently than men or women. Those who were unclassified also endorsed non-binary experiences more frequently than cisgender or transgender people.

### **Conclusion**

In general, the results of these two studies support our hypothesis that the multiple-choice gender identity question (alongside the transgender identity question) would have adequate

substantive validity, generalizability, and external validity for use with transgender and cisgender populations, and for use with non-binary/agender people. Consistent with the bandwidth-fidelity tradeoff, this measure does a good job of predicting outcomes specifically related to gender *identity* (e.g. self-identification, self-investment in gender groups), but is a poorer predictor of more distantly related concepts (e.g. pronoun usage, clothing choices). However, this same tradeoff is present for the binary gender question (and all the other gender questions we tested). As described in chapters 1 and 2, a single question cannot capture precise information about all gender-related constructs.

We also tested a measure of transgender identity alongside the multiple-choice gender identity question. Initially, we hoped to use these results to show that gender identity is the same construct in transgender and cisgender people, but this approach did not take into account that transgender identity *intersects* with gender identity in a way that influences the meaning of that identity (much as socioeconomic status, age, race, and sexual orientation do). For example, transgender people reported greater distress in response to identity denial than cisgender people, presumably because their identities are routinely invalidated in the course of daily life. As all people have multiple identities which influence one another, the question of whether gender identity is identical for cisgender and transgender people now seems easily answered; the subjective experience of gender identity differs for everyone. The *construct* of gender identity, however, is theoretical, and must be measured the same way across all participants in order to enable meaningful comparison. Transgender identity should be measured alongside gender identity when it is relevant to the topic of research, because it *can* intersect with gender identity in a way that may influence a variable of interest. However, it need not always be included, any more than age or SES are always included in every study.

Perhaps the greatest improvement the multiple-choice gender identity question offers over the binary gender question is its explicit inclusion of non-binary and agender identities. These results suggest that non-binary and agender identities predict unique patterns of self-definition and self-investment in gender groups, emotions toward gender groups, and pronoun usage, distinct from those of men or women. Gender identity measures which include options for non-binary and agender people are not merely better-liked, but also produce more valid data by separating groups which systematically differ from the commonly recognized “man” and “woman.”

## Summary of Chapters 3-5

Chapters 3 through 5 present the results of two studies (four samples) designed to determine which of the previously recommended transgender and non-binary inclusive gender measures would be most appropriate for general use. To determine this, I assessed the simplicity of administration and analysis of each question (Chapter 3) and its level of acceptability to participants (Chapter 4), and identified the best-performing measure. I then provided evidence for the validity of that measure, the multiple-choice gender identity measure, in Chapter 5.

### **Chapter 3: Simplicity of Inclusive Gender Identity and Transgender Status Measures**

Chapter 3 is concerned with how easy it is to aggregate data from each measure, and outlines detailed plans for coding and condensing the sometimes large amounts of information yielded by the various gender measures. That is, **I aimed to identify the simplest measure which is still inclusive of transgender and non-binary people, and to clearly describe a plan for using the data which it produces.** I also explored the content validity, general usability, and reliability of each measure. The consistency between different gender measures was generally quite high. Major differences between question types were logistical (e.g. ease of coding) rather than quantitative. From a standpoint of simplicity, I recommend the binary + other or multiple-choice gender identity question, along with the transgender identity or transition history question (if transgender status is needed).

#### **Gender Identity Measures.**

The multiple-choice gender measure identified more diversity than the binary gender question, as was logically and theoretically anticipated. In Study 1, 34% of participants were classified as a gender other than man or woman by the multiple-choice question, and in Study 2, 7% were (Study 2 had a greater proportion of undergraduate students than Study 1).

Correspondence between the binary + other question and the multiple-choice gender measure was quite high (85% Study 1, 95% Study 2), and correspondence between the free-text and multiple-choice gender question was similarly high (88% Study 1, 96% Study 2). As this overlap between questions is substantial, it does not particularly recommend one question over another.

Interestingly, across both studies, responses to the free-text question could be *meaningfully* categorized (i.e. into men, women, non-binary, or agender responses) somewhat more often than responses to the multiple-choice question. This was true despite the fact that the multiple-choice question includes an option to write in a free-text response, which was coded and categorized by hand using the same coding scheme. For applications where gender identity is the primary interest and research support is available to code responses, the free-text question may be appropriate. However, the multiple-choice question is much less logistically difficult to analyze. The number of responses requiring hand-coding is much smaller, significantly reducing research investment necessary to interpret the data. In addition, the free-text question requires researchers to determine how to group participants' responses (a constantly changing problem) rather than allowing them to group themselves. Ironically, this apparently maximally inclusive question in practice gives participants very little agency over how their data are ultimately interpreted by researchers.

While there was substantial overlap between the coding categories developed from participants' free-text responses and the options available in the multiple-choice gender identity question, I identified several necessary alterations for content validity based on Study 1 results that were implemented in Study 2. Most notably, I changed the "genderqueer" response option to "non-binary" and added a response option for "agender."



Pilot results from two alternative types of gender self-categorization measures suggest further directions for improving the content validity beyond the Study 2 gender identity measures. Contrary to Tate et al.'s hypothesis that non-binary people identify mostly as "both men and women" while agender people mostly identify as "neither men nor women," I found that most non-binary *and* agender people identified as neither men nor women on a measure created by Joel et al. (2014). Another important dimension for future research concerns the effects of social recognition on identification with various gender identities. I adapted a biracial identity question created by Rockquemore and Arend (2002) to measure gender identity (especially among non-binary people). I found great variability in the gender categories which participants perceive as available to them (e.g. the emergence of "non-binary" as an umbrella term between Study 1 and Study 2), the categories they feel others assign to them (e.g. being perceived as binary in daily life), and the strength of their identification with any of these available categories. Further research is needed to determine how non-binary and agender people conceptualize their identities. One thought-provoking approach to this question, Sexual Configurations Theory, was proposed by van Anders in 2015 and is still being explored. While this conceptualization is not yet suited to general use in surveys, it may yield insights that can eventually improve the content validity of more conventional gender measures.

### **Transgender Status Measures.**

I also conducted a supplemental analysis of transgender status measures. Conceptually, transition history, transgender identity, and assigned sex at birth are different constructs, so I did not expect perfect correspondence between them. Nonetheless, when comparing across the three transgender status measures, 83% of participants were classified the same way (as Not Transgender, Transgender, or Unclassified), and most of the remaining participants were

classified the same way by two out of three methods. There was not perfect overlap between any of the methods. These results alone suggest that the measures are quite similar in the total proportion of participants they identify as transgender.

However, the sex assigned at birth question is more complex to interpret than the other measures as it must be compared to a gender identity measure to generate transgender status. It also has an important usability concern with respect to non-binary and agender participants. In a typical coding scheme for this measure, *all* non-binary/agender people are either transgender (if transgender is defined as having any gender identity different from sex assigned at birth), or *no* non-binary or agender people are transgender (if transgender is defined to include only people with binary gender identities). In this case, sex assigned at birth provides no additional information about transgender status beyond what is obvious from gender identity.

For general use, the transgender identity question and transition history question are equally easy to administer and analyze because they can be used as-is without the recoding which is necessary for the sex assigned at birth question. The transgender identity question has our final recommendation because it has greater face validity for concepts related to identity (which is of greater relevance for social psychology).

#### **Chapter 4: Acceptability and Identity Safety of Gender Measures**

Chapter 4 considered whether gender measures may act as cues to identity safety for transgender and non-binary participants, and how these cues may differ across various contexts in which researchers might use gender measures. My goal was to **identify the gender measures that are most acceptable to participants of varying gender identities and transgender statuses, and that provide the greatest sense of identity safety across contexts.** Considering these results alongside my recommendations for simplicity of analysis, I recommend that

researchers use the multiple-choice gender identity question for general purposes, alongside the transgender identity measure if information about transgender status is needed.

### **What question do participants prefer?**

The short free-text question was best-liked and most identity safe regardless of the context in which it appeared, participant gender, or transgender status. The multiple-choice gender identity question was next most liked/identity safe, and the binary gender question was disliked and viewed as identity-unsafe. Both transgender and cisgender participants use inclusive gender questions to infer researchers' views of gender minorities, and have improved perceptions of researchers and increased comfort with research when seeing inclusive questions. This suggests that using inclusive gender measures may help to increase participation and retention of participants, both transgender and cisgender.

### **Which questions should researchers use?**

The short free-text question was best-liked and most identity-safe, but as described in Chapter 3, the complexity of analyzing its results make it unwieldy for general use. I recommend the multiple-choice gender question for most research purposes, as it is better liked by participants across contexts in both studies than the binary + other question. In Study 2, I experimentally demonstrated that participants who saw the multiple-choice question had higher expectations of belonging, higher comfort with research, and lower negative emotions toward researchers than those who saw the binary. In contrast, those who saw the binary + other question did not differ from those who saw the binary question on those measures.

If information about transgender status is needed, I suggest using the multiple-choice + transgender identity question, as it was best-liked and most identity-safe among transgender and non-binary participants.

When considering all participants, the multiple-choice + assigned sex at birth and multiple-choice + transgender identity measures were rated approximately equally. However, the higher ratings of the transgender identity question among gender minorities coupled with the increased logistical simplicity of analyzing its results (especially for non-binary/agender people) means that it has my recommendation. However, for medical contexts where physical bodies are of greater relevance, the assigned sex at birth measure is still a viable option, unlike the multiple-choice + transition history measure (which was so strongly disliked in Study 1 that I dropped it from Study 2).

## **Chapter 5: Validating Inclusive Gender Identity Measures**

Chapter 5 presents evidence about the substantive validity, general usability, and external validity of the gender measure which had the best combination of simplicity of analysis and acceptability/identity safety for participants: the multiple-choice gender identity measure, alongside the transgender identity measure. **Broadly, I wanted to provide empirical evidence that this gender measure captures the information it promises and is not measuring conceptually unrelated constructs.** In general, this was true; the multiple-choice gender identity measure is a good predictor of constructs that are very closely related to gender identity (e.g. self-investment in gender groups) and a less precise predictor of more distantly related concepts (e.g. clothing choice).

These analyses aimed to show that the multiple-choice gender measure predicts constructs closely related to/moderated by gender identity, such as self-investment and self-definition in gender groups (substantive validity) as well as gender-identity related but more distant constructs (external validity), like pronoun usage and clothing choices. These hypotheses were generally well-supported.

I also hoped to show that gender identity is the same construct for transgender and non-binary/agender participants as it is for cisgender and binary participants to demonstrate the general usability of the measure. That is, I expected participants across gender identities and transgender statuses to nonetheless show similar affiliation with their gender identity groups (as measured by gender centrality, gender identification, and emotions toward their gender group). This hypothesis was less well supported. Gender centrality and positive emotions toward one's gender group were highest for transgender people, while gender identification was highest for cisgender men and women. On further consideration, these results seem to reflect the ways that transgender identity *intersects* with gender identity to influence the meaning of that identity; for example, considering the great barriers many transgender people face in achieving social recognition of their genders, it makes sense that those gender identities would have greater salience and centrality. This further underscores the importance of measuring transgender identity alongside gender identity when gender is a topic of central interest, because transgender identity can interact with gender identity in distinct ways.

### **Conclusion**

The binary and binary + other gender measures and the transition history transgender status measure should not be used, because their very low level of acceptability and identity safety poses a risk to the recruitment and retention of participants (especially gender minorities). I recommend the multiple-choice gender identity measure for general use based on its high sensitivity and simple analytic requirements, high acceptability and identity safety across participants and contexts, and ability to successfully predict theoretically related constructs with a moderate level of precision. When a measure of transgender status is needed, I recommend the transgender identity measure on similar grounds.

## **Chapter 6: Effects of Gender Identity Measures on Attitudes toward Gender Minorities**

Until this point, I have considered how the use of various transgender-inclusive gender identity measures will affect researchers (in terms of each measure's simplicity of use and validity) and gender minorities (in terms of the acceptability of each measure). In this chapter, I consider the effects that transgender-inclusive gender identity measures may have on the general population's attitudes and beliefs about gender minorities.

Very few researchers have seriously considered the effects of particular gender question formats on the general population. The work which does exist has mostly asked whether questions about transgender identity might make cisgender people uncomfortable (much in the way that researchers wondered whether asking questions about sexual orientation might make heterosexual people uncomfortable). Although most cisgender participants seem to have no difficulty with or objection to the two-question method, a small minority of these participants may be uncomfortable or offended by questions which imply that sex and gender can be distinguished, that gender is not binary, and/or that transgender people exist (Conron et al., 2008; Redford & Van Wagenen, 2012; Balarajan et al., 2011; Lombardi et al., 2013).

Borrowing from the literature on racial demographic questions, however, we may speculate that any new gender question which becomes widely adopted (and is thus viewed by large sections of the population) will have a different set of effects on society at large; that is, it will subtly influence the lay public's views of gender. Lee, Wilton and Kwan (2014) find that demographic questions which require individuals to identify with only one ethnicity ("check one" as opposed to "check all that apply" questions) serve as a subtle cue to racial essentialism. This increased racial essentialism, in turn, led racial minorities (Asian-Americans) to show reduced interest in cultural diversity events, but also led European-Americans to be less

interested in intergroup contact with Asian-Americans. Lee et al. (2014) further conclude that essentialist cues generally lead participants to expect less intergroup contact.

Generalizing from this study to the context of gender, we may therefore predict that gender demographic questions which require a forced choice between categories will cue essentialist beliefs. That is, after seeing a single-select gender identity question, essentialist views may become more salient (especially for people who already had relatively essentialist views), resulting in greater acceptance of inequality, increased prejudice, and decreased interest in intergroup contact.

If these relationships are eventually demonstrated, various other social psychological findings about essentialism may be tested in the context of gender. For example, racial essentialism (conceiving of race as biologically based) leads both racial minorities and racial majorities to be more accepting of racial inequality and less interested in intergroup contact (e.g. Williams & Eberhardt, 2008). Of greatest relevance, Ching and Xu (2017) found that experimentally increased bioessentialist beliefs (specifically that the gender binary is biological) also increased negative stereotyping of transgender people, and transprejudice. If forced-choice gender questions tend to cue (or even increase) gender essentialism to the same degree that forced-choice race questions increase racial essentialism, researchers' choice of gender question formats may affect intergroup relations and the experience of prejudice directly.

Encountering an inclusive gender identity measure may also produce reactance (Brehm, 1966) to the extent that the measure is seen as threatening the values of a meaningful group to which participants belong (de Lemus, Bukowski, Spears, & Telga, 2015). People who already essentialize gender may perceive these inclusive gender identity measures as threatening to their values around gender. As suggested by the Sacred Value Protection Model (Tetlock et al., 2000),

for people who view the idea that gender is binary as a “sacred value,” merely considering a gender measure which undermines that value will produce a feeling of violation and subjective contamination (Hauptert, 2015). This threat may manifest in moral outrage toward and denigration of the perceived source of the moral threat (e.g. gender minorities, the researchers) and a reaffirmation of the threatened moral values (e.g. the gender binary, normative gender roles).

Building on the design used by Lee et al. (2014) to study the effects of racial demographic questions on racial essentialism, we designed an experiment to determine the effect of differently-worded gender demographic questions on gender essentialism and prejudice toward gender minorities (among other outcomes). In line with the findings of Lee et al. (2014), we hypothesize that for participants with gender essentialist views, seeing single-select gender identity measures (relative to seeing questions that permit participants to choose more than one option) will cue essentialism and/or produce reactance, resulting in more negative attitudes toward gender minorities. We also tentatively hypothesize that single-select gender identity measures may increase participants' endorsement of gender essentialist beliefs.

## **Method**

### **Study 2 - Pilot Test of Novel Gender Essentialism Scale**

Haslam, Rothschild, and Ernst (2000) defined essentialism as a belief that a members of a social category share an underlying natural, innate “essence.” We wanted to measure these beliefs about gender, but found that the only available published measure of gender essentialism, the Gender Theory Questionnaire (Coleman & Hong, 2008) only assessed participant’s beliefs about the extent and origin of gender differences (i.e. whether they are social or biological). While those concepts are related to essentialism as defined by Haslam et al. (2000), several other



concepts are not assessed, including Discreteness, Uniformity, Necessity, and Exclusivity. To ensure that we assessed the full range of essentialism-related concepts and to increase our ability to compare gender essentialism to other forms of essentialism, we created our own measure by adapting and adding to the Race Conceptions Scale developed by Williams and Eberhardt (2008). This measure corresponds much more closely to the nine elements of essentialism initially proposed by Haslam et al. (2000).

The following scale was included in Study 2, which was conducted in Fall 2016 (see Chapters 3-5 for details about participants and the remainder of the methods). Items are grouped here by the element of essentialism (Haslam et al., 2000) which they are measuring, but were randomized when presented to participants. Participants indicated their level of agreement with each of the following 28 items on a 7 point Likert scale (1 = Strongly disagree, 7 = Strongly agree). Items with an asterisk were reverse-coded.

Discreteness (2 items)

- *Men and women are not fundamentally different.\**
- *It's easy to tell what gender someone is by looking at them.*

Uniformity (3 items)

- *People who are of different genders may look quite similar to each other.\**
- *Generally speaking, [two men/two women] will always look more similar to each other than a man and a woman ever would.*

Informativeness (3 items)

- *Knowing someone's gender doesn't tell you very much about them.\**
- *Knowing someone's gender tells you a lot about them.*
- *I can make a lot of judgments about a person just by knowing their gender.*

Naturalness (2 items)

- *It's natural to notice a person's gender.*
- *Young children probably learn about which people fall into which gender categories automatically, without much help from adults.*

Immutability (3 items)

- *A person's gender is fixed at birth.*
- *Whether a person is a man or a woman is pretty much set early on in childhood.*

- *A person's gender can change over the course of their life.\**

#### Stability (6 items)

- *There's agreement across cultures about which gender groups people fall into.*
- *The political climate can dictate how someone's gender is categorized.\**
- *The same gender categories have pretty much always existed.*
- *Gender categories haven't always existed in the world.\**
- *In 200 years, society will use basically the same gender categories.*
- *How a person's gender is defined depends on the social context.\**

#### Inherence (3 items)

- *Genders are innate, genetically-based tendencies.*
- *It's impossible to determine a person's gender by examining their DNA.\**
- *Gender is caused by biological factors.*

#### Necessity (4 items)

- *There is no defining characteristic that is necessary to be a [man/woman].\**
- *[Men/Women] have a necessary or defining characteristic, without which they would not be [men/women].*

#### Exclusivity (2 items)

- *It's possible to be a full member of more than one gender.\**
- *It's possible to be both a man and a woman.\**

A scree test showed a clear, sharp drop after the first factor, supporting our hypothesis that the scale is measuring a single factor. Scores had high reliability ( $\alpha=.94$ ) and were aggregated into a single composite score ( $M = 3.8$ ,  $SD = 1.1$ ).

### **Study 3 – How Gender Identity Measures Affect Prejudice Toward Gender Minorities**

Participants were randomly assigned to answer one of four gender questions (some with single-select options, and some with multiple-select options and/or free text boxes) followed by my novel measure of gender essentialism. They then completed several previously validated scales in two blocks with measures presented in randomized order within each block. The first block contained the Homophobia Scale (Wright, Adams, and Bernat, 1999), Rosenberg Self-Esteem Scale (Rosenberg, 1965), Genderism and Transphobia Scale (Nagoshi et al., 2008), Transphobia Scale (Hill & Willoughby, 2005), Ambivalent Sexism Scale (Glick and Fiske,

1996), and the Implicit Person Theory Scale (Dweck et al., 1995) in a random order. The second block contained the Gender Theory Questionnaire, (Coleman & Hong, 2008), Right-Wing Authoritarianism (Altemeyer, 1981), General Essentialism Scale (Bastian & Haslam, 2006), the Social Desirability Short Form (Strahan & Gerbasi, 1972), and the Need for Closure Short Form (Roets & Ban Hiel, 2011). Finally, participants completed standard demographic questions and were debriefed about the purpose of the study.

### **Participants.**

In the spring of 2017, I recruited a convenience sample of undergraduate psychology majors from the psychology department participant pool at a large public university in the Midwest United States. There was a 94% completion rate; 468 participants consented and 442 completed the entire survey. Participants received course credit for participating.

Ages ranged from 16-39 ( $M = 19.6$ ,  $SD = 1.94$ ). With respect to gender identity, participants ( $N = 440$ ) were 36% men ( $N = 159$ ), 63% women ( $N = 276$ ), 0.4% genderfluid ( $N=2$ ), and 0.2% each non-binary, do not know, and choose not to answer (each  $N = 1$ ). With respect to racial/ethnic identity, participants ( $N=440$ ) were 8% Black ( $N=36$ ), 0.4% Pacific Islander ( $N=2$ ), 2.9% Indian Subcontinent ( $N=13$ ), 0.9% Middle Eastern ( $N=4$ ), 13.6% East Asian ( $N=60$ ), 1.5% Southeast Asian ( $N=7$ ), 73.4% White ( $N=323$ ), 4.5% Latinx ( $N=20$ ), and 1.1% Native American ( $N=5$ ).<sup>20</sup> One participant wrote in another racial identity (“Armenian”). Politically, participants ( $N = 437$ ) were 43.2% Democrat ( $N = 189$ ), 23.3% Republican ( $N = 102$ ), 27.2% Independent ( $N = 119$ ), and 6.1% Other party ( $N = 27$ ). Examples of participants write-in responses were “None,” “I don’t think I have any party thoughts,” “I’m foreign, from a socialist country, would choose democrat,” “Libertarian,” and “Justice Democrat.”

---

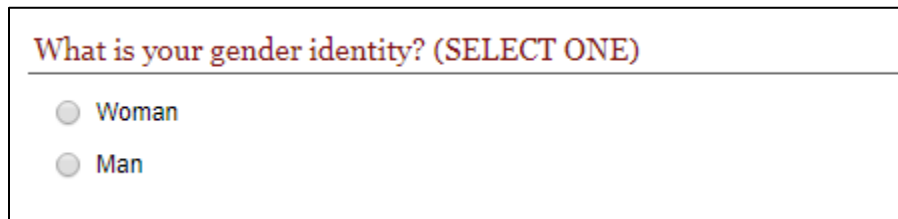
<sup>20</sup> Participants could select more than one race/ethnicity, so percentages sum to more than 100%.

## Materials.

### *Gender Demographic Questions.*

Participants were randomly assigned to answer one of four gender questions. For the binary and binary + other questions, the first two response options were presented in a random order, and for the multiple-choice questions the first three response options were randomized.

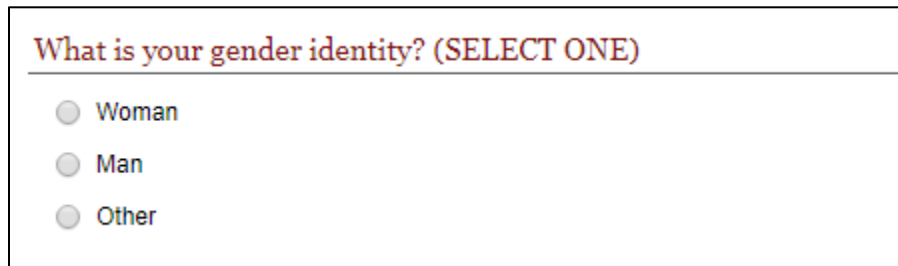
The binary question (Figure 1) was included as a baseline, as it is both the most commonly used gender question and one that does not challenge the dominant cultural conception of gender as binary.



A screenshot of a survey question. The question text is "What is your gender identity? (SELECT ONE)" in red. Below the question, there are two radio button options: "Woman" and "Man".

*Figure 1. Single-select binary gender identity question.*

The single-select binary + other question (Figure 2) was included for comparison to the baseline binary gender question. It acknowledges the possibility of genders other than man and woman, but because it does not name these possibilities and instead literally “others” them, I expect that participants’ levels of gender essentialism will not differ between the binary question and the binary + other question.

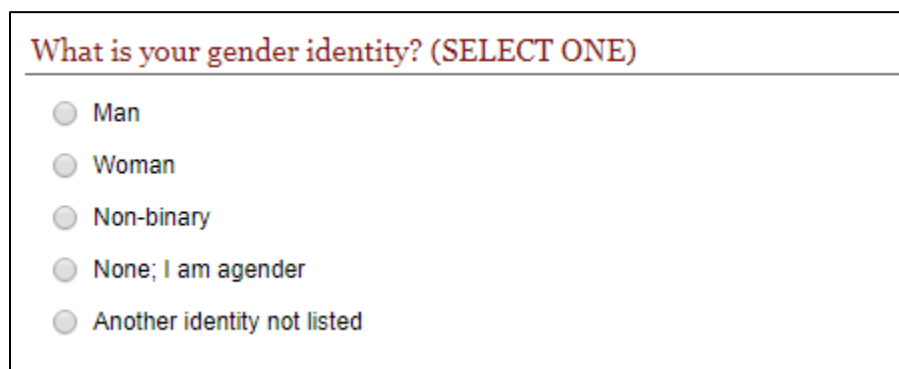


A screenshot of a survey question. The question text is "What is your gender identity? (SELECT ONE)" in red. Below the question, there are three radio button options: "Woman", "Man", and "Other".

*Figure 2. Single-select binary + other gender identity question.*

In contrast, the single-select multiple-choice gender question (Figure 3) both acknowledges the existence of multiple genders and names them explicitly. However, it still

forces participants to select only one of these options, reinforcing the essentialist idea that these gender identities are mutually exclusive. If naming genders other than man and woman is sufficient to reduce essentialism, the essentialism of participants who see this question should differ from those who see the binary and binary + other questions. However, if allowing the selection of multiple genders is necessary to reduce essentialism, participants who see this question should have essentialism scores similar to those who saw the binary or binary + other questions.

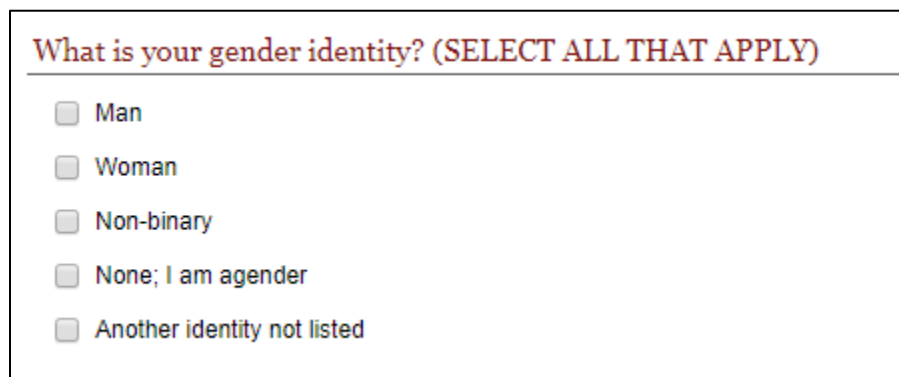


What is your gender identity? (SELECT ONE)

- Man
- Woman
- Non-binary
- None; I am agender
- Another identity not listed

Figure 3. Single-select multiple-choice gender identity question.

Finally, the multi-select multiple-choice gender question (Figure 4) was included to test the hypothesis that allowing the selection of more than one gender undermines the essentialist idea that gender identities are mutually exclusive. If this is so, participants who see this gender question should have lower essentialism scores than participants who see any of the other three gender questions.



What is your gender identity? (SELECT ALL THAT APPLY)

- Man
- Woman
- Non-binary
- None; I am agender
- Another identity not listed

Figure 4. Multi-select multiple-choice gender identity question.

***Moderator Variables.***

*Novel Gender Essentialism Measure.*

I revised my novel gender essentialism measure to include only the 19 items with the highest item-score correlation in the Fall 2016 surveys. All items used a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree). Reverse-coded items are marked with an asterisk.

1. *Generally speaking, two men will always look more similar to each other than a man and a woman ever would.*
2. *Generally speaking, two women will always look more similar to each other than a man and a woman ever would.*
3. *People who are of different genders may look quite similar to each other.\**
4. *Knowing someone's gender tells you a lot about them.*
5. *It's natural to notice a person's gender.*
6. *The same gender categories have pretty much always existed.*
7. *Gender categories haven't always existed in the world.\**
8. *It's easy to tell what gender someone is by looking at them.*
9. *Whether a person is a man or a woman is pretty much set early on in childhood.*
10. *A person's gender can change over the course of their life.\**
11. *A person's gender is fixed at birth.*
12. *There's agreement across cultures about which gender groups people fall into.*
13. *Gender is caused by biological factors.*
14. *Women have a necessary or defining characteristic, without which they would not be women.*
15. *There is no defining characteristic that is necessary to be a woman.\**
16. *Men have a necessary or defining characteristic, without which they would not be men.*
17. *There is no defining characteristic that is necessary to be a man.\**
18. *Men and women are not fundamentally different.\**
19. *It's possible to be a full member of more than one gender.\**

The scree plot showed a clear, sharp drop after the first factor, supporting the hypothesis that the scale is unidimensional. Therefore, we aggregated these 19 items into a single composite score ( $\alpha = .88$ ,  $M = 4.2$ ,  $SD = 0.96$ ).

*Gender Theory Questionnaire (Coleman & Hong, 2008).*

Participants indicated their agreement with each of 11 items measuring gender essentialism on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include “When men and women differ in some way, it is likely that the difference is due to biological factors.” and “It is hard if not impossible to change the innate dispositions of a person’s gender.” Although a scree test suggested that this scale has two factors, we averaged all item scores into a single composite score ( $\alpha = .77$ ,  $M = 3.9$ ,  $SD = .83$ ) as suggested by the authors. Higher scores represent greater gender essentialism.

***Outcome Variables.***

*Genderism and Transphobia Scale (Nagoshi et al., 2008).*

Participants indicated their agreement with each of nine items measuring negative attitudes toward transgender people on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include “I don’t like it when someone is flirting with me, and I can’t tell if they are a man or a woman.” and “I think there is something wrong with a person who says that they are neither a man nor a woman.” I averaged item scores into a single composite score ( $\alpha = .91$ ,  $M = 3.4$ ,  $SD = 1.3$ ); higher scores indicate more negative attitudes toward transgender people.

*Transphobia Scale (Hill & Willoughby, 2005).*

Participants indicated their agreement with each of 32 items measuring negative attitudes toward transgender people on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include “Sex change operations are morally wrong.” and “Feminine men make me feel uncomfortable.” Although Hill and Willoughby divide their scale into three subscales, the scree plot showed strong evidence of unidimensionality, so I averaged item scores

into a single composite score ( $\alpha = .97$ ,  $M = 2.6$ ,  $SD = 1.1$ ); higher scores indicate more negative attitudes toward transgender people.

### ***Additional Variables.***

We included seven other scales in this survey as part of a project to validate the novel gender essentialism scale. These were the Homophobia Scale (Wright, Adams, and Bernat, 1999), Rosenberg Self-Esteem Scale (Rosenberg, 1965), Social Desirability Short Form (Strahan & Gerbasi, 1972), Need for Closure Short Form (Roets & Ban Hiel, 2011), General Essentialism Scale (Bastian & Haslam, 2006), Right-Wing Authoritarianism (Altemeyer, 1981), Ambivalent Sexism Scale (Glick and Fiske, 1996), and Implicit Person Theory Scale (Dweck et al., 1995). At the end of the survey, participants completed six demographic questions about gender identity, sexual orientation, race/ethnicity, age, political affiliation, and religion. A brief description of each of these variables is in Appendix D. We do not report further analyses for these measures.

## **Results**

In line with Lee et al. (2014), we predicted that seeing single-select gender identity measures (relative to seeing questions that permit participants to choose more than one option) would cue essentialism and/or produce reactance, resulting in more negative attitudes toward gender minorities and possibly increasing participants' endorsement of gender essentialist beliefs.

### **Regressions by Gender Demographic Question Condition**

I ran multiple regressions predicting Transphobia (Hill & Willoughby, 2005) and Genderism and Transphobia (Nagoshi et al., 2008) from gender demographic question condition (Binary, Binary + Other, Single-Select Multiple-Choice, or Multi-Select Multiple-Choice), with the Binary question as the comparison group. Neither the regression for Transphobia ( $F(3, 412)$ )



= .45,  $p = 0.71$ ) nor the regression for Genderism and Transphobia ( $F(3, 436) = 0.56, p = 0.63$ ) was significant.

I also ran multiple regressions predicting the novel gender essentialism measure (Hauptert, 2016) and the Gender Theory Questionnaire (Coleman and Hong, 2008) from gender demographic question condition (Binary, Binary + Other, Single-Select Multiple-Choice, or Multi-Select Multiple-Choice), with the Binary question as the comparison group. Neither the regression for novel gender essentialism ( $F(3, 419) = 1.91, p = 0.12$ ) nor the regression for the Gender Theory Questionnaire ( $F(3, 427) = 1.07, p = 0.36$ ) was significant.

Thus, exposure to different versions of the gender identity question did not have statistically reliable direct effects on either transphobia or gender essentialism (two somewhat different measures of each).

### **Interaction Analyses**

We also conducted analyses predicting transphobia from a combination of gender question condition, novel gender essentialism (Hauptert, 2016), and the Gender Theory Questionnaire (Coleman & Hong, 2008). These analyses correspond to the findings of Lee et al. (2014), who found effects of question type only among participants who were already high in essentialism.

#### **Gender Question Condition and Gender Essentialism (Hauptert, 2016).**

There was no significant interaction of novel gender essentialism and gender question condition for Transphobia ( $F(3, 382) = 0.63, p = 0.59$ ) or Genderism and Transphobia ( $F(3, 407) = 0.21, p = 0.88$ ). There was a main effect of novel gender essentialism on Transphobia ( $F(1, 385) = 165.05, p < .001$ ); higher novel gender essentialism predicted higher Transphobia ( $t = 12.8, p < .001$ ). Likewise, there was a main effect of novel gender essentialism on Genderism and

Transphobia ( $F(1, 410) = 415.58, p < .001$ ); higher scores on the novel gender essentialism predicted higher Genderism and Transphobia ( $t = 20.38, p < .001$ ).

**Gender Question Condition and Gender Theory Questionnaire (Coleman & Hong, 2008).**

There was no significant interaction of Gender Theory Questionnaire and gender question condition for Transphobia ( $F(3, 394) = 1.06, p = 0.36$ ) or Genderism and Transphobia ( $F(3, 418) = 0.34, p = 0.79$ ).

There was a main effect of Gender Theory Questionnaire on Transphobia ( $F(1, 397) = 135.96, p < .001$ ); higher Gender Theory Questionnaire predicted higher Transphobia ( $t = 11.6, p < .001$ ). Likewise, there was a main effect of Gender Theory Questionnaire on Genderism and Transphobia ( $F(1, 421) = 248.73, p < .001$ ); higher scores on the Gender Theory Questionnaire predicted higher Genderism and Transphobia ( $t = 15.77, p < .001$ ).

### **Discussion**

Our initial hypotheses were not supported. None of the gender question conditions directly affected gender essentialism or transphobia, and there was no interaction between gender question condition and gender essentialism on the measures of transphobia. Thus, our study did not find effects of question condition either overall or specifically among the participants who scored highest on gender essentialism.

These results are not consistent with those reported by Lee, Wilton, and Kwan (2014) in their research on racial demographic questions. Our outcome variables differed from theirs, in that we assessed gender essentialism and transphobia directly, while they measured implicit person theory and interest in intergroup contact. They prioritized unobtrusive measurement and concealment of their research goals, while we prioritized comprehensive measurement. It seems

unlikely that Lee et al.'s relatively unobtrusive measurement approach accounts for our differing results; given our inclusion of a wide variety of other types of measures in the survey and the ubiquity of gender identity questions in surveys, it seems unlikely that participants were engaged in hypothesis-guessing.

An alternative explanation is that single-select gender identity questions are simply not as strong a cue to gender essentialism as single-select race demographic questions are to racial essentialism. Indeed, the vast majority of gender identity questions that people encounter in daily life are single-select (see Chapter 2), such that many participants may not even notice them. However, the results of my Study 2, especially those reported in Chapter 4, show that merely seeing a binary gender identity question *does* produce a significant effect on perceptions of researchers, suggesting that participants have at least some awareness of the type of gender identity question they see.

Another explanation for our failure to obtain results paralleling Lee, Wilton, and Kwan (2014) in the context of gender demographic questions is that gender may already be a chronically accessible construct for the vast majority of people in a way that race is not for white people. Thus, the relative effect of any particular cue to gender essentialism might be reduced (e.g. Bargh, Lombardi, & Higgins, 1988).

Future research could include manipulation checks for attention to the wording of each measure and evaluation of each measure to examine the possibility that an effect would emerge with more attentive participants (although given that Lee, Wilton, and Kwan (2014) propose that this effect is based on a subtle cue to racial essentialism, such a change to the experimental design would undermine the generalizability of the results to a non-experimental setting). Researchers could also attempt to unobtrusively measure and control for the chronic accessibility

of gender constructs to each participant (although if gender is a chronically accessible construct for most people, this may have limited effect on the results). Finally, reducing the number of measures included in each study would reduce survey fatigue and might increase the chances of detecting a small effect.

With respect to our goal of recommending a gender identity measure for general use, however, these results suggest that any effect of gender identity question type on gender essentialism or transphobia is likely to be quite limited in practice, and so that concern should not be a deciding factor when choosing what method to use.

## Chapter 7: Including the Gender Spectrum in Measures of Sexual Orientation

Males do not represent two discrete populations, heterosexual and homosexual. The world is not to be divided into sheep and goats. Not all things are black nor all things white. It is a fundamental of taxonomy that nature rarely deals with discrete categories. Only the human mind invents categories and tries to force facts into separated pigeon-holes. The living world is a continuum in each and every one of its aspects. The sooner we learn this concerning human sexual behavior, the sooner we shall reach a sound understanding of the realities of sex. (Kinsey, 1948, p. 639)

While classifications of sexual orientation have existed since at least the time of Plato (Reeve, 2006), scientific interest in sexual orientation dates back only to 1860 and the work of Karl Ulrichs, whose classification scheme roughly mapped on to the heterosexual/bisexual/homosexual distinction still widely used today (Sell, 1997). This is not coincidental; according to Sell (1997), Ulrichs' views influenced nearly all the early researchers of sexual orientation, including Kraft-Ebbing and Hirschfeld, who through Freud and Jung spread these ideas to the rest of the psychological community. Writing nearly a century after Ulrichs, Alfred Kinsey's great contribution to the scientific view of sexual orientation was to describe it as a continuum ranging from heterosexual to homosexual rather than a set of discrete categories, operationalized in his famous Kinsey Scale. Yet despite his assertion that "the living world is a continuum in *each and every one* of its aspects" (p. 639, emphasis added), Kinsey's work on sexual orientation (like that of Ulrichs and even Plato before him) reinforced the traditional Western view of gender as a set of two discrete categories.

Indeed, the assumption that the world is divided into men and women is the premise from which all modern measures of sexual orientation begin. To indicate my position on the Kinsey Scale, for example, I must first be able to identify my own gender as male or female, after which I must identify the proportions of men and women among my preferred partners. That genders other than male and female might exist, or that the single term "gender" might conceal a

constellation of constructs ranging from chromosomes to clothing style, does not feature in the Kinsey Scale. Even more complex multidimensional sexual orientation measures such as Shively and DeCecco's sexual orientation scale (1977) or Klein's Sexual Orientation Grid (1985) make no mention of either the ill-defined construct of gender or the assumption that gender is binary in their critiques of the Kinsey Scale. Nor have new conceptualizations or measures of sexual orientation been forthcoming in the decades since the 1980s (though for notable exceptions see e.g. Diamond, 2003 and van Anders, 2015). On the contrary, Vrangalova and Savin-Williams (2012) note that the three categories of heterosexual, bisexual, and homosexual "have become so culturally and politically entrenched in contemporary societies that they have achieved the status of 'natural kinds,' that is, naturally occurring rather than socially constructed distinctions" (p. 85). Clearly, scientists and society alike have settled on a sexual orientation measure, one with the gender binary at its core.

Yet for several decades now, researchers from a variety of disciplines have slowly problematized the view of gender as a dichotomy between male and female. As early as 1955, John Money and his colleagues were distinguishing between sex (biological factors) and gender (social factors), and by 1972, they described chromosomal, gonadal, and hormonal sexes as well as gender identities and roles (Muehlenhard & Peterson, 2011). Subsequent research on intersex conditions (e.g. Fausto-Sterling, 2000; Consortium, 2006) has made clear that chromosomes, gonads, hormones, genitals, and secondary sex characteristics can each take multiple forms, and while certain combinations are more common, the form of any one cannot perfectly predict the forms of the others. Likewise, many transgender theorists have critiqued the view of gender identities and roles as binary and categorical (e.g. Bornstein, 1996; Butler, 2006; Califia, 1997), a view which is supported by recent theory and empirical research in psychology (Tate et al.,

2014; Joel et al., 2014). Converging evidence would seem to suggest that gender is vastly more complex than the man/woman dichotomy that underlies standard sexual orientation measures, but researchers have yet to fully acknowledge, much less address, this essential problem with the construct.

### **Who is Harmed by the Status Quo?**

Retaining the standard sexual orientation question is detrimental to people with marginalized sexual and gender identities and impedes progress toward a scientific understanding of sexuality and its relationship to gender. In particular, the experiences of people with non-binary gender identities cannot be fit into the standard categories, while those of transgender people are reshaped to better fit the standard model (and often inconsistently at that). This systematic inattention to these most marginalized groups has already negatively affected their health and social outcomes, and is likely to continue to do so until researchers alter their practices to include (or perhaps even focus on) these communities. From a methodological standpoint, too, imprecise definitions and systematic inattention to data which contradict commonly-held views are antithetical to the scientific method and the pursuit of greater understanding about the world around us. Taken together, these are compelling reasons to revise our measures of sexual orientation and think broadly about new ways to conceptualize these constructs.

#### **Transgender people with binary gender identities.**

Researchers attempting to collect and report the sexual orientations of transgender people were among the first scientists to be confronted with the ambiguity of traditional sexual orientation measures. Among the most immediate questions to be answered was this: how should a person's sexual orientation be reported when their assigned sex and gender identity differ?

Sadly, instead of using this opportunity to rethink the assumptions underlying sexual orientation measures, early psychologists simply defined away the problem, choosing to bend transgender participants' experiences to fit the standard heterosexual/homosexual model rather than acknowledging their embodiment of the model's limitations. For example, in the field of clinical psychology/psychiatry, "nearly all investigators have referenced the sexual orientation of transsexuals to birth sex" (Lawrence, 2010, p. 515) in preference to transgender people's self-definitions, presentations, or even their genitals and hormonal makeup. Such an approach cannot even be called bio-essentialist. Far from asserting a particular scientifically-based location or criteria for gender, the decision to privilege "birth sex" merely reifies the (culturally constructed) gender binary by means of an appeal to the equally culturally constructed concept of sex *assigned* at birth.

In a similar move to de-emphasize the disruptive potential of considering transgender experiences, the Institute of Medicine's recent report on the health of LGBT individuals (2011) states that "whether a sexual act or romantic attraction is characterized as homosexual or heterosexual depends on *the biological sex* of the individuals involved, relative to each other" (p. 27, emphasis added). On the next pages, the authors "clarify" their previous definition by stating that "sexual orientation refers to an enduring pattern of sexual or romantic activity with men, women, *transgender persons*, or some combination of these groups" and acknowledging that some people may "have a specific attraction toward *transgender persons*" (p. 28, emphasis added). This wording casts transgender people as possible objects of desire while denying their existence as agents capable of desiring. Thus, even as trans identities are erased by describing biological sex as the only appropriate site by which to determine the "homosexuality" or "heterosexuality" of an act, trans identities are recognized as possible targets of desire.



Setting aside this simultaneous erasure and eroticization of trans identities, as well as the authors' assumption that transgender people cannot be men or women, their statements do nonetheless suggest a recognition that gender is not a unitary construct and that genders other than male and female can exist. While this realization could have led to a critique of commonly-held ideas about sexual orientation, the authors chose instead to work within the existing gendered framework for sexual orientation, opting to simplify the picture of sexual orientation rather than acknowledging its probable complexity.

Yet this dogged adherence to traditional views of sexual orientation fails to simplify the situation – rather, it renders suspect any data collected using such measures, as we have no satisfying definition of the criteria used to assess sexual orientation and no guarantee that the researchers and participants were using the same criteria. For example, consider a transgender woman who has sex with women. If the researchers ask her for her sex assigned at birth, she may report “woman,” even if her original birth certificate stated otherwise (e.g. Balarajan et al., 2011). If the researchers ask her for her sexual orientation, she will likely report that she is a lesbian, which the researchers may or may not “translate” into an assumption of heterosexuality under their assigned-sex-based model. Even if the researchers ask her whether she has sex with men or women, she may (quite reasonably) report the gender identities of her partners rather than their sex assigned at birth, but researchers may assume that her partners are cisgender.

This approach to research with transgender people betrays a deep-seated distrust of transgender people's lived, self-reported experience, a belief that their identities are not merely non-normative but also non-real. For scholars to hold such prejudices against their participants is unconscionable on ethical grounds alone, but it also renders their data next to useless. Given all of the definitional uncertainties described previously, interpreting other researchers' data on

transgender people's sexual orientations is difficult if not impossible. Social science, and indeed society at large, lacks a consistent, strictly applied operational definition of gender on which to anchor sexual orientation, and without such a definition, researchers cannot hope to conduct meta-analyses or even draw simple conclusions from one another's work.

### **People with non-binary gender identities.**

As we have seen from the experiences of binary-identified transgender people, it can be difficult to fit sexual desires, attractions, and behaviors into a researcher-approved sexual orientation category even among people who identify as men and women. These difficulties are only compounded further for people with non-binary gender identities and people who are attracted to people with non-binary genders. Current measures define sexual orientation in the relationship between your gender and the gender(s) of your partners, which makes a certain amount of sense in a world with two and only two genders; if the researchers know your gender and your sexual orientation, they know what gender(s) your partners will have.

If we acknowledge the existence of more than two genders, however, knowing your gender and your sexual orientation won't always let researchers know what gender(s) your partners will have. A heterosexual genderqueer person may be attracted to (any one of) men, women, agender people, bigender people, etc., while a bisexual person of any gender may be attracted to any combination of two or more genders. This imprecision means that many non-binary individuals (and their partners) will be unable to fully convey their sexual identities to researchers, while leaving researchers unaware that they are missing any information.

Furthermore, the intelligibility of these categories seems to depend entirely on how literally non-binary people are willing to interpret the labels. Even the theoretically straightforward case where a person is attracted other people of their own gender seems complex

for participants in practice. For example, consider the perspective of this participant quoted in Galupo et al. (2014):

“As a person who is gender queer and who prefers the same in partners, I have a hard time figuring out if I am homosexual or not! It depends on the solidity of your gender category which I don’t have” (p. 441).

This quote points to two important considerations. One is that participants are generally motivated to give researchers the information they’re asking for, and will consider the context of the entire survey in order to do so (Schwartz, 1999). Common sexual orientation questions thus put non-binary participants in an awkward situation, unable to answer in a way that serves their needs (represents their self-identity) or their perceptions of the researcher’s needs (to accurately categorize them). The second, related point is that individuals with non-binary gender identities are likely to interpret sexual orientation questions in the context of the type and quality of gender questions included in the survey (see Chapter 4 for evidence about identity threat from question wordings). As the person above indicates, the choice of sexual orientation category for a non-binary person depends in large part on the solidity they perceive in the researchers’ conceptions of gender. A researcher who asks a very clear, non-binary inclusive gender question may be perceived as having more “solid” gender categories, which may prompt participants like the one above to identify themselves differently than they otherwise would.

These considerations become particularly relevant as minority-inclusive gender questions are used more frequently; while gender questions can signal an awareness of non-binary gender identities, sexual orientation questions have not kept pace. Even surveys conducted by and for transgender and gender non-conforming people have been reluctant to rethink the status quo of sexual orientation measures. For example, the National Transgender Discrimination Survey, the largest and most comprehensive studies of transgender experiences to date, included four

separate, innovative questions assessing various aspects of gender identity and presentation and was in nearly all respects highly inclusive of gender minorities (Grant et al., 2011). Yet they included only a single question assessing sexual orientation, and that question still adheres closely to the norm despite adding “queer” and “asexual” categories and a free-text box. Even looking back on this after the fact, the authors of this report argued that “there were very few true ‘others’ that didn’t approximate the concepts that were listed” (p. 201) and contended that the free-text option should be dropped.

69. What is your sexual orientation? <input type="radio"/> Gay/Lesbian/Same-gender attraction <input type="radio"/> Bisexual <input type="radio"/> Queer <input type="radio"/> Heterosexual <input type="radio"/> Asexual <input type="radio"/> Other, please specify _____
---

*Figure 1. Sexual orientation question from the National Transgender Discrimination Survey (Grant et al., 2011).*

Of course, sexual orientation was not the primary research interest for these authors, and they were dealing with significant constraints on survey length, participant interest, and resources for coding responses. Yet these obstacles exist for the vast majority of social science researchers, and it is somewhat disappointing that even a team of researchers (many of whom are themselves transgender) focused solely on the experiences of transgender people would not attempt to adjust this particular question to better capture the experiences of people with non-binary genders.

No doubt some will argue that labels like “queer” are inclusive of people with non-binary genders, which to a limited extent is true. However useful the word may be from a political and social standpoint, the term “queer” is too broad and too difficult to operationally define to be particularly useful for most empirical research, as it is used differently by different people and

can encompass nearly all non-heterosexual identities. Thus, the term “queer” may serve as little more than a glorified “other” box, concealing an array of different experiences which might provide insight into a variety of research problems. For that matter, participants who select the “other” box are also likely to have sexual identities which break the mold, and there may be quite a lot of them. In a recent review of data from several nationally representative U.S. samples, Carpenter (2013) reports that between 0.2% and 4.0% of the population selects the “Other” or “Not Sure” option on sexual orientation questions if one is provided. In comparison, he finds that between 1% and 2.3% of the general population identifies as gay or lesbian.

In addition to their relative uselessness to researchers, survey options like “other” and “queer” may be emotionally fraught for some participants, discouraging their use. Participants in a UK cognitive interview about the design of minority-inclusive gender questions argued that the term “other” has “connotations of being strange and is offensive” (Balarajan et al. 2011, p. 42), a sentiment which is probably shared by US participants indicating their sexual orientations. Those over 35 are more likely to regard the word “queer” as a homophobic slur than a reclaimed term of pride and self-definition. In one notable example, a 66-year-old gay man complained to the Colorado Attorney General after seeing the word used in the (voluntary) demographic section of a Colorado College job application (Cotton, 2013).

As we see here, researchers can create a cycle of invisibility merely by using the current sexual orientation framework, even with minor adjustments like the addition of options like “other,” “queer,” and/or “please specify” options. Although non-binary people cannot express their sexual identities in the current framework, the question structure gives them few ways to indicate their dissatisfaction. As described previously, the “other” and “queer” catch-alls are little help to researchers, and the inclusion of free-text boxes is unlikely to solve the problem either.

So long as the coders who analyze these data conceptualize sexual orientation as heterosexual/bisexual/homosexual, even responses which clearly deviate from the expected pattern will be interpreted through its lens and repackaged as “approximating” the recognized categories. Researchers using the standard sexual orientation measures will have no evidence that anyone is dissatisfied, and without evidence of a problem, future researchers will imitate past researchers.

Despite this tendency to perpetuate invisibility, some researchers *have* asked more nuanced questions about participants’ perceptions of standard sexual orientation questions, and their data corroborate many of these more theoretical objections to their use. Galupo et al. (2014) conducted an online survey of 285 convenience-sampled participants who identified as sexual minorities (i.e. nonheterosexual people). Participants completed both the Kinsey Scale and the Klein Sexual Orientation Grid and gave free-text responses to the following question: “In what ways did this scale capture or fail to capture your sexuality?” Thematic analyses revealed several main themes in these free-text responses, including the idea that sexual orientation is about self-identification rather than behavior<sup>21</sup>; that sexual orientations and identities are influenced by sociocultural factors; and that identities can be fluid.

Most importantly for our purposes, several transgender participants and participants with non-binary genders commented that sexual orientation measures do not fully capture their sexual identities, both because they are based in a gender binary and because they cannot account for fluidity in gender and sexual identities. They further noted feeling uncertainty about defining sexual orientation on the basis of sex; as one participant said;

---

<sup>21</sup> This point is frequently acknowledged by researchers creating more modern sexual orientation measures, but is not always preserved when results of those measures are used (e.g. in health care settings).

“this scale asks me to talk about the sex characteristics of the people I am attracted to, not their gender identity or presentation, which is much more of a factor in my own attraction than sex characteristics” (p. 449).

It seems that, when given the chance to express their opinions about standard sexual orientation questions, many people (especially transgender people or those with non-binary gender identities) find the current sexual orientation questions painfully limiting and inadequate.

### **Partners of transgender people and/or those with non-binary gender identities.**

Because sexual orientation is an inherently relational construct, sexual orientation measures that misrepresent and overlook transgender and non-binary people will do the same to their partners. Partners of people with non-binary gender identities have no way to express this on current sexual orientation measures, while the partners of binary-identified transgender people cannot be sure that researchers will correctly or consistently identify their partners' gender. In either case, the limitations of current sexual orientation measure may result in a miscategorization of their sexual orientations by proxy. As one queer trans man quoted in Garofalo (2014) wondered,

“Is my partner ‘really gay’ as some have argued he’s not, because he’s with a female-bodied transman? Am I incidentally hetero if I’m with a pre-transition transguy who’s physically female?” (p. 449).

This quote points to yet another concerning consequence of the psychological tradition of classifying transgender people’s sexual orientations on the basis of sex assigned at birth. Conceptualizing sexual orientation in this way reifies the fear among cisgender men that being attracted to or having sex with a transgender woman means they’re “actually gay”; after all, that’s exactly how most psychologists and the Institute of Medicine currently classify such activities. As fears of identity misclassification (being misperceived as homosexual) seem to underlie both homophobia and transphobia in cisgender men (Nagoshi et al., 2008), this particular classification scheme exacerbates the already high risk of discrimination and violence

transgender women regularly face (Grant et al., 2011). Even when not contributing to outright violence against gender minorities, researchers who use standard sexual orientation measures prevent their participants from expressing desire for transgender and non-binary people, which perpetuates the simultaneous invisibility and intense stigmatization of these desires.

### **What are the Practical Problems with Current Sexual Orientation Measures?**

By misinterpreting and ignoring the sexual identities and experiences of transgender and non-binary people, current sexual orientation measures may compromise the health and social outcomes of gender minorities. This happens most frequently when clinicians, practitioners, or researchers use sexual orientation as a proxy for other information of interest (e.g. fertility, STI risk) without acknowledging or accounting for the error this introduces. If researchers interpret participants' responses in ways that are inconsistent with the participants' lived experiences, researchers risk drawing factually inaccurate conclusions from their data (as well as alienating and marginalizing their patients/participants). This inaccurate data may in turn be used to support inappropriate care for individuals as well as counterproductive health and social policies, harming gender minorities and society more generally. In the following section, I will explore one major area in which sexual orientation measures are inappropriately used with negative consequences for gender minorities: the provision of health care.

Several major medical and public health organizations recognize sexual and gender minorities as among the highest-risk, most marginalized populations from a public health standpoint (e.g. Institute of Medicine, 2011; Department of Health and Human Services, 2013; American Psychological Association, 2008), and all have identified inadequate demographic questions as a barrier to improving care. The scientific community has failed to consistently and accurately collect information about the most marginalized groups, despite the fact that the



members of these communities are most in need of research and intervention. When viewed from this perspective, the frequently-raised objection that more specific or complex sexual orientation questions might be inconvenient or uncomfortable for “the majority” of patients (i.e. cisgender heterosexual people; Redford & Van Wagenen, 2008) seems a small price to pay.

Sexual orientation information is perhaps most ubiquitously (and, I would argue, inappropriately) used in two primary health care circumstances; when providing contraceptives, and when preventing and treating sexually transmitted infections. In the following sections I explore the reasons why sexual orientation questions are inappropriate for these circumstances and suggest practical alternatives.

#### **Birth control and fertility.**

Although clinicians routinely rely on sexual orientation questions to determine which patients should receive birth control, sexual orientation is a poor proxy for fertility or the likelihood of pregnancy. Consider, for example, the number of potentially flawed assumptions being made by any clinician who offers a sexually-active heterosexual woman a prescription for birth control pills and assumes she must want to get pregnant when she turns it down (cf. Morrison & Dinkel, 2012).

This presumes first that all women are cisgender and have the (functional) reproductive organs necessary to *carry* a pregnancy, which isn't the case for some intersex women, trans women, or cis women with any of various medical conditions resulting in infertility. Of course, trans women may also want birth control if they are capable of impregnating a partner, and providers must be willing and able to offer condoms and vasectomies to these women rather than assuming that gender identity indicates the appropriate type of birth control.

Next, the clinician presumes that a heterosexual woman's partner is always a cisgender man with the (functional) reproductive organs necessary to impregnate her. As noted previously, trans women may wish to prevent themselves from impregnating a partner, while even heterosexual cis women capable of becoming pregnant may be partnered with infertile cis men or with trans men, or people with non-binary gender identities who do not have the necessary reproductive organs to impregnate them. Alternatively, they may partner with other cis women while still identifying as heterosexual (e.g. Vrangalova & Savin-Williams, 2012); in each of these cases, birth control would be unnecessary.

Finally, the clinician assumes that all heterosexual women who say they are sexually active mean that they are engaging in types of sexual activity capable of making them pregnant. However, even fertile cisgender women who are sexually active with fertile cisgender men are not necessarily engaging in types of sex that can result in pregnancy. Recent research indicates a great deal of variability in perceptions of the meaning of the phrase "had sex." For example, oral and anal sex as well as mutual masturbation seem to constitute "sexual activity" to a significant portion of the population (Sanders et al., 2010) but virtually never result in pregnancy. Conversely, 10.9% of respondents felt that penile-vaginal intercourse did not constitute "having sex" if it was not accompanied by ejaculation, and 5.2% felt that penile-vaginal intercourse was not "having sex" even if it did include ejaculation (Sanders et al., 2010), despite the potential risk of pregnancy with these sexual activities.

Given the degree of inaccuracy introduced by each of the previous assumptions, clinicians who use sexual orientation as a proxy for birth control needs are doing a great disservice to their patients. Simply asking patients whether they would like to learn about birth control options as a starting point would likely be equally effective, regardless of their gender,

fertility status, sexual orientation, sexual partners, or sexual activities. In addition, such questions might be more likely to result in educating cisgender men and people who *plan* to have types of sex that may cause pregnancy but haven't done so yet. Using this more inclusive initial question gives patients the agency to describe their own contraceptive needs proactively, rather than passively reacting to a series of assumptions from their medical provider. After the conversation is started, medical practitioners can follow up with more specific questions (e.g. Are you interested in hormonal or barrier or surgical methods? How long-lasting do you want the contraceptive technique to be?) as they usually would.

### **Sexually transmitted infections.**

In addition to contraception counseling, medical practitioners and public health workers frequently use sexual orientation (and to a lesser extent the nominally-better-defined relationship between one's gender and the gender of one's sexual partners) as a proxy for a person's risk of contracting a sexually transmitted infection (STI). However, these questions virtually never operationally define the "sex" or "gender" used to anchor their assessment of sexual orientation or sexual behavior. They also fail to take into consideration the potential differences in definition used by practitioners and their patients; much like psychologists and their trans patients, public health workers may think "sex" means sex assigned at birth, while their respondents are more likely to think "sex" means current physical status and/or gender identity. When assessing STI risk, these differences are crucially important.

Blood donation policies are one example. Even after recent proposed reforms take effect, men who have had sex with men (MSM) in the past year will still be considered ineligible to donate blood. While this deferral is controversial (and even opposed by the Red Cross), the FDA argues that the policy reflects appropriate caution considering the high rates of HIV among MSM

(FDA, 2015). This assessment, although crude, seems to have more grounding in evidence; the FDA reports that 61% of all new HIV infections in the United States in 2010 were among men who have sex with men, although these people make up only about 2% of the population (FDA, 2015). However, the prevalence of HIV among transgender women (28%) may be as high as that among even the most marginalized group of men who have sex with men (young urban black men), yet no blood donation screening questions even attempt to identify transgender women (Ard & Makadon, 2012).

While the FDA may regard transgender women as “men” by virtue of their sex assigned at birth, they certainly do not define themselves as such; trans women are therefore unlikely to indicate that they are “men who have sex with men” when asked during donor screening. By failing to operationally define “sex” or “gender” for in their donor regulations (or making any other provisions for identifying gender minorities), the FDA hypocritically permits a readily identifiable high-risk population (transgender women) to donate blood. Of course, an ideal deferral policy would be based as closely as possible on information about the actual risk behavior of interest (e.g. penile-anal intercourse, multiple sex partners, coerced sex), rather than using a person’s gender as a proxy. Unfortunately, the FDA is unlikely to revise its position until research yields evidence of the efficacy of alternate screening questions, and the inconsistency with which it applies its own policies in the meantime is deeply concerning. Equally concerning is the lack of research attention being directed toward better understanding the sexual lives of non-heterosexuals and transgender people (IoM, 2011; Ard & Makadon, 2012), who are surely among the most marginalized and at-risk populations of Americans from a public health standpoint.

### **Necessary Features of an Inclusive Sexual Orientation Measure**

In a first attempt to address this long-standing problem, we designed a set of sexual orientation questions designed to be inclusive of transgender and non-binary identities as well as asexual identities. Following previous research on the usability of standard sexual orientation questions for transgender, non-binary, and asexual individuals (e.g. Bogaert, 2013; Galupo et al., 2014; Garofalo, 2014; van Anders, 2015), we identified the following features as necessary for an inclusive sexual orientation measure:

**1. Gender must not be viewed as binary.**

As a bare minimum requirement, an inclusive sexual orientation question must provide a way for participants to indicate an attraction to or behavior with transgender people and/or non-binary people. Commonly used sexual orientation measures like the Kinsey Scale (Kinsey, Pomeroy, & Martin, 1948) and the Klein Grid (Klein, 1985) are anchored by binary gender categories (i.e. men and women), and are thus inherently non-inclusive of transgender and non-binary identities. In addition, they generally provide no integrated way to indicate a lack of attraction to or behavior with anyone, making them non-inclusive for asexual identities. These are not merely issues for inclusion, but also for data validity (see Chapter 1).

**2. Participants' genders must be measured separately from genders they are attracted to.**

Sexual identity categories like “heterosexual,” “homosexual,” and “bisexual” are fundamentally statements about the relationship between one’s own gender and the genders of one’s partners, and are predicated on the assumption that there are only two genders. As such, they lose clarity and meaning when they encounter non-binary identities. For example, as discussed earlier, it is conceptually unclear whether a man who is attracted to women and genderqueer people should label himself “heterosexual” (as he is attracted to people of genders

other than his own) or “bisexual” (as he is attracted to people of more than one gender), and equally unclear what sexual identity labels (if any) are appropriate for non-binary people.

Researchers and participants also disagree (among themselves and with each other) about whether assigned sex at birth or gender identity serve as the conceptual underpinning of gender in sexual orientation. That is, when assigned sex and gender identity differ, which should one reference when describing one’s sexual orientation? In clinical psychology “nearly all investigators have referenced the sexual orientation of transsexuals to birth sex” (Lawrence, 2010, p. 515), and the Institute of Medicine (2011) argues that “whether a sexual act or romantic attraction is characterized as homosexual or heterosexual depends on *the biological sex* of the individuals involved, relative to each other” (p. 27, emphasis added). However, previous research with transgender people and their partners makes clear that many of them declare their sexual identities on the basis of gender identity instead (e.g. Balarajan et al., 2011; Galupo, 2014).

Both of these factors introduce a great deal of noise into the meaning of sexual identity labels, and make predicting the target genders of a participants’ attractions and behaviors from their sexual identities an extremely confusing prospect. The binary assumptions inherent in common sexual orientation questions are also a source of great dissatisfaction for participants who have relationships with trans and non-binary people or who are trans or non-binary themselves (Galupo, 2014), potentially undermining their engagement/participation in health care and research settings which use these questions (Meier & Labuski, 2013).

### **3. Participants must be able to indicate attraction to multiple genders, one, or none in an integrated way.**

As described in the previous paragraphs, categorical identity-based measures make it impossible to tell which genders a person is attracted to once the assumption of binary genders

has been lifted. Likewise, measures like the Kinsey Scale do not provide an integrated way to indicate attraction to genders other than “men” and “women” or a lack of attraction to anyone. To reduce social desirability effects, an inclusive measure must provide for all of these possibilities without implying that one scenario is “typical.”

**4. The components of sexual orientation (identity, attraction, and behavior) must be measured separately.**

There is general consensus among sexuality researchers that sexual orientation is a multidimensional construct. While researchers may differ in the exact constructs and definitions they use, a distinction between identity, attraction, and behavior is very common (e.g. Rosario & Schrimshaw, 2014; Bancroft, 2009). Researchers of bisexuality (e.g. Klein, 1985; Weinberg, Williams, & Pryor, 1994) and asexuality (e.g. Bogaert, 2013) have also observed that people may have different sexual and romantic attractions. Therefore, to remain up-to-date with modern conceptions of sexual orientation and appropriately capture the nuance of participants’ experiences, these constructs should be measured separately.

## **General Method**

### **Overview**

While a variety of measures have the capacity to meet the requirements for transgender, non-binary, and asexual-inclusive sexual orientation questions we have just described, we chose to focus on sexual orientation questions suitable for inclusion in standard demographic sections of surveys and screening forms. Rather than using a Likert-type scale as done in the Kinsey Scale, we opted for a multi-select model. Participants were shown a multi-select list of genders, as well as a “Nobody” option and a write-in choice, and asked to indicate which genders of people they were attracted to or had ever had sex/romantic relationships with. They were asked

about their sexual attraction, sexual behavior, romantic attraction, and romantic behavior separately.

We then pilot-tested these questions with undergraduate students for three semesters to determine their usability as well as their divergence from a standard single-item identity-based sexual orientation measure. We hypothesized that most participants would be able and willing to answer the questions, that identity, romantic attraction, sexual attraction, romantic behavior, and sexual behavior would emerge as distinct (but related) constructs, and that participants with a variety of sexual identities would report attraction to and behavior with transgender and non-binary people.

## **Participants**

Participants were undergraduate students at a large Midwestern university. Our questions were included in the mass testing survey conducted each semester, which contains demographic questions and pre-screening measures for various experiments being conducted in the psychology department. At the start of each semester, all students in several introductory-level psychology classes are strongly encouraged to complete the survey, and they receive course credit for doing so.

## **Materials**

### **Binary gender question and traditional sexual orientation question.**

A binary gender question (What is your gender? *Male/Female*) and a traditional sexual orientation question (What is your sexual orientation? *Heterosexual/Straight; Homosexual/Gay/Lesbian; Bisexual/Pansexual; Other (Asexual, Questioning)*) were included as part of the standard demographics in every mass testing survey.

### **Other gender questions.**



As part of another study we were conducting, participants were asked up to four transgender inclusive gender questions, including a pair of questions about gender identity and sex assigned at birth, and a series of questions about the strength of their identification with various gender identities. For the sake of space, we have not yet considered the sexual orientation questions in light of participants' responses to these other gender questions, but we plan to do so in future. However, it is possible that the presence of these questions immediately before our new sexual orientation questions shaped participant responding.

### **New sexual orientation questions.**

As described in the introduction, we designed a series of four questions to assess two major dimensions of sexual orientation with relevance to gender - *attraction* and *behavior*. We further subdivided these dimensions into *sexual* and *romantic* forms. These questions were as follows:

- *What genders of people are you sexually attracted to?*
- *What genders of people are you romantically attracted to?*
- *What genders of people have you had any kind of sex with?*
- *What genders of people have you had romantic relationships with?*

All four questions had the same basic response options (participants could select multiple):

- *Cisgender men*
- *Cisgender women*
- *Transgender men*
- *Transgender women*
- *Genderqueer people or people with other non-binary genders*
- *I am not (romantically/sexually) attracted to anyone. / I have not had (any kind of sex/romantic relationships) with anyone.*
- *Other [free text box]*

### **Procedure**

All survey questions were included in the mass testing survey each semester, and presented alongside a variety of measures being collected for other psychology experiments. The

mass testing survey was administered through Qualtrics, and participants completed it online at a time and location of their choice. The binary gender and traditional sexual orientation questions were always asked before the trans-inclusive gender questions and the new sexual orientation questions, but the number and kind of questions presented in between differed from semester to semester.

### **Hypotheses and Plan for Analysis**

To reduce participants' responses to the new sexual orientation questions, we created 20 new categories to represent common response patterns. These were as follows:

- Participants who selected no options (1 category)
- Participants who selected only one option (7 categories, one for each choice)
- Participants who selected cisgender men and one other choice (6 categories, one for each combination)
- Participants who selected cisgender women and one other choice (5 categories, one for each combination)
- Participants who selected three or more choices (1 category)

Participants were placed into one of the twenty categories for each of the four new sexual orientation questions. As a comparison point, we categorized participants according to their responses to the binary gender and traditional sexual orientation question, yielding eight possible categories (heterosexual man, heterosexual woman, gay man, lesbian woman, bisexual man, bisexual woman, other man, other woman).

We defined and tested our hypotheses as follows.

#### **Hypothesis 1: Participants will be able and willing to answer these new questions.**

First, we expected that most participants would answer the questions rather than leaving them completely blank. Participants who had missing data in otherwise complete surveys (i.e. who viewed all the new sexual orientation questions but did not answer them) were categorized as unwilling to answer these questions.

Next, we expected that most participants would not feel the need to write in responses in the provided “Other” option, as we hoped that our response options already covered most possibilities that participants would want to report. We also recognized that participants might use the free text box to express dissatisfaction with the design or wording of the question. Whatever the cause, we considered use of the “Other” option to be another indicator of participants’ unwillingness and/or inability to answer the new questions.

Finally, we expected that the majority of participants’ responses to these new questions would resemble their answers to the traditional sexual orientation question. While we argue that identity, attraction, and behavior are different constructs (i.e. they are theoretically distinct, and do not always occur in any particular combination), previous research and theory still suggests that these constructs are consistent for a majority of people, in part because people often base their identities on their own attractions and behavior. If responses to our new sexual orientation questions diverged too dramatically from responses to the traditional sexual orientation question (for example, if most heterosexual men indicated sexual attraction to cisgender men), we would have reason to doubt the validity of our new questions.

**Hypothesis 2: Sexual identity is a distinct construct from attraction and behavior.**

Following previous research (cf. Rosario & Schrimshaw, 2014), we expect that sexual identity as measured by the traditional sexual orientation question is a distinct construct from sexual attraction and behavior. Thus, we predict that for some participants, their reported sexual attractions and behaviors will differ from what might be expected based on their sexual identity. However, as described in Hypothesis 1, this does not mean that we expect all participants to show different responses to the traditional sexual orientation question and the new ones. Because

the concept of “difference” between these constructs can be defined in a variety of ways, we tested this hypothesis in two ways.

We first used a liberal definition of “difference,” working from the strong version of assumptions we imagined researchers might make about participants based on their sexual identities. For our liberal definition, we presumed that sexual identity implied the following:

- *Heterosexual identity* implies sexual attraction and behavior with cisgender men or nobody (for women) or cisgender women or nobody only (for men).
- *Gay/lesbian identity* implies sexual attraction and behavior with cisgender men or nobody (for men) or cisgender women or nobody (for women).
- *Bisexual/pansexual identity* implies sexual attraction and behavior with cisgender men and/or women or nobody.

We did not make specific predictions about the attractions or behaviors of people who selected the “Other” option, as we thought most researchers would likewise not make assumptions about the attractions or behaviors of this category.

Next, we used a conservative definition of “difference,” using a weaker version of the assumptions we imagined researchers would make. For our conservative definition, we presumed the following:

- *Heterosexual identity* implies no sexual attraction or behavior with cisgender women (for women) or cisgender men (for men).
- *Gay/lesbian identity* implies no sexual attraction or behavior with cisgender women (for men) or cisgender men (for women).

We did not make specific predictions about the attractions or behaviors of people who identified as bisexual/pansexual or who selected the “Other” option.

**Hypotheses 3 and 4: Attraction and behavior are distinct constructs, and they each have distinct romantic and sexual forms.**

We expected that in a substantial proportion of cases, participants' responses to the attraction and behavior questions would differ from each other, and that their responses to the romantic and sexual questions would also differ. As with Hypothesis 2, we tested this hypothesis with two different definitions of "difference."

We first used a liberal definition of difference. If participants were grouped into different categories for ANY of the four new sexual orientation questions, their answers were considered "not consistent." We next used a more conservative definition of difference. We omitted inconsistencies because of the "Nobody" category, reasoning that at least some researchers who use the sexual orientation question would be aware that people do not always act on their attractions. We also noted that many of those who selected "Other" simply added more detail about their preferences for an already-listed gender, rather than listing another gender. Thus, we did not consider responses "inconsistent" if the only difference between categories came from selecting "Nobody" or from writing in an explanation.

**Hypothesis 5: Sexual orientation is not exclusively toward cisgender people or people with binary genders.**

We expected to find that some proportion of participants are sexually and/or romantically attracted to transgender and non-binary people. Likewise, we expected that some proportion of participants would have sexual and/or romantic relationships with transgender and non-binary people. We regarded this hypothesis as supported if any participants indicated sexual or romantic attractions or behaviors with transgender men, transgender women, or non-binary people.

**Hypothesis 6: Sexual/romantic attraction and behavior with transgender/non-binary people will occur across sexual identities.**

We expected that people with a variety of sexual identities would experience sexual/romantic attraction and sexual/romantic behavior with transgender and non-binary people. As described in the introduction (e.g. Lawrence, 2010; Galupo, 2014), researchers and participants vary in their categorizations of people who are attracted to binary transgender people. Some argue that assigned sex at birth is the relevant category (such that e.g. cis men attracted to trans women are homosexual), while others, including most trans people, argue that gender identity is the relevant category of comparison (such that e.g. cis men attracted to trans women are heterosexual). In addition, while terms like “pansexual” are generally interpreted as signaling attraction to/behavior with people with non-binary gender identities, we expected that at least some people who experience such attraction and behavior would use other identity terms (e.g. using “heterosexual” to mean “attraction to any gender other than my own”). Thus, due to a lack of cultural consensus around the identity implications of attraction to and behavior with binary and non-binary trans people, we expected to see these attractions and behavior among people from a variety of identity categories.

### **Study 1: Fall 2014**

The participants, materials, and procedure in the Fall 2014 semester were identical to those described in the general method, with one exception. Due to a programming error, the question “*What genders of people are you sexually attracted to?*” had the response option “*I am not romantically attracted to anyone.*” and the question “*What genders of people have you have romantic relationships with?*” had the response option “*I have not had any kind of sex with anyone.*” A total of 2850 participants completed the entire survey. According to the binary gender and traditional sexual orientation question, the sample was 38% heterosexual men, 56%

heterosexual women, 0.6% lesbian women, 1.1% gay men, 0.7% bisexual men, 1.5% bisexual women, 0.2% “other” men, and 0.6% “other” women.

**Results and Discussion**

**Hypothesis 1.**

Hypothesis 1 was generally supported; most participants seemed able and willing to answer the questions. Only 20 responses of 2850 (0.7%) were missing, which compares favorably to the 49 missing responses (1.7%) on the binary gender question. Some participants (between 2.8% and 3.6% depending on the question) did feel the need to write in responses, suggesting that the usability of the question can still be improved. We also noted that a small minority of participants (generally 2-3) selected the “Nobody” option and another option on each question, which does not make sense and suggests that some participants may have difficulty understanding the questions. Overall, though, most participants’ responses to the new sexual orientation question were congruent with their responses to the traditional sexual orientation question (see Hypothesis 2 for more detail), suggesting that most participants understood the questions and were able to answer them.

**Hypothesis 2.**

Hypothesis 2 was supported, consistent with previous research. Using a liberal interpretation of their sexual identities, the sexual attractions of 9.8% of participants were misclassified, and the sexual behavior of 6.5% was misclassified. Even with an extremely conservative interpretation of their sexual identities, the sexual attractions of 0.9% of participants were misclassified, and the sexual behavior of 0.3% was misclassified.

*Table 1. Inconsistent responses on sexual identity and sexual attraction/behavior – Fall 2014.*

	<b>Liberal definition</b>				<b>Conservative definition</b>			
	Attraction		Behavior		Attraction		Behavior	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%

<b>All participants</b>	279	9.8	185	6.5	25	0.9	10	0.3
<b>Heterosexual Men</b>	99	3.5	75	2.6	9	0.3	4	0.1
<b>Heterosexual Women</b>	151	5.3	84	2.9	11	0.4	2	0.1
<b>Gay Men</b>	5	0.2	10	0.4	-	-	1	<0.1
<b>Lesbian Women</b>	8	0.3	9	0.3	5	0.2	3	0.1
<b>Bisexual Men</b>	5	0.2	4	0.1				
<b>Bisexual Women</b>	11	0.4	3	0.1				

### Hypotheses 3 and 4.

Hypotheses 3 and 4 were supported in a general sense. Between 30.5 % and 7.2% of participants had inconsistent responses on at least one of the four new sexual orientation questions (depending on definition), and there were participants with inconsistent answers across the four most plausible combinations of sexual and romantic attraction and behavior. However, the numbers of people with inconsistent responses dropped dramatically between the liberal and conservative definition, suggesting that most (though not all) of this inconsistency derives from people who have sexual/romantic attractions but have not acted on them.

*Table 2. Inconsistent responses on sexual/romantic attraction/behavior – Fall 2014.*

	<b>Liberal definition</b>		<b>Conservative definition</b>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
<b>Any inconsistent responses</b>	868	30.5	206	7.2
<b>Romantic and Sexual Attraction Inconsistent</b>	208	7.2	140	4.9
<b>Romantic and Sexual Behavior Inconsistent</b>	500	17.5	80	2.8
<b>Romantic Attraction and Behavior Inconsistent</b>	368	12.9	100	3.5
<b>Sexual Attraction and Behavior Inconsistent</b>	792	27.8	149	5.2

### Hypothesis 5.

Hypothesis 5 was supported; 3.6% of participants reported sexual/romantic attraction toward or sexual/romantic behavior with trans or non-binary people. Attraction (2.6%) was more common than behavior (1.3%), and proportions of attraction and behavior were roughly equivalent between trans men, trans women, and non-binary people.

*Table 3. Sexual/romantic attraction/behavior with non-cisgender people – Fall 2014.*

	<b>Any non-cis</b>	<b>Trans men</b>	<b>Trans women</b>	<b>Non-binary</b>



	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>Any attraction/behavior</b>	103	3.6	54	1.9	58	2.0	55	1.9
<b>Romantic attraction</b>	73	2.6	40	1.4	43	1.5	42	1.5
<b>Sexual attraction</b>	73	2.6	30	1.1	42	1.5	33	1.2
<b>Romantic behavior</b>	36	1.3	10	0.4	15	0.5	16	0.6
<b>Sexual behavior</b>	40	1.4	15	0.5	16	0.6	20	0.7

### Hypothesis 6.

Hypothesis 6 was supported; a small percentage of participants from almost all sexual identities reported sexual or romantic attraction or behavior with a trans or non-binary person. However, this was clearly more common among GLB participants (especially those who identified as bisexual/pansexual) than among heterosexuals, considering the much larger proportion of heterosexual people in the sample and the population. There was also no clear relationship between sexual identity and relative frequency of attraction or behavior with trans men, trans women, or non-binary people. That is, people from each identity category appeared to be attracted to/having relationships with trans people with a variety of gender identities and assigned sexes.

*Table 4. Any sexual/romantic attraction/behavior with non-cis people by sexual identity – Fall 2014.*

	<b>Any non-cis</b>		<b>Trans men</b>		<b>Trans women</b>		<b>Non-binary</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>Heterosexual men</b>	45	1.6	17	0.6	29	1.0	16	0.6
<b>Heterosexual women</b>	22	0.8	13	0.5	6	0.2	10	0.4
<b>Gay men</b>	5	0.2	3	0.1	1	<0.1	3	0.1
<b>Lesbian women</b>	3	0.1	2	0.1	2	0.1	2	0.1
<b>Bisexual men</b>	6	0.2	5	0.2	4	0.1	5	0.2
<b>Bisexual women</b>	15	0.5	11	0.4	11	0.4	14	0.5
<b>“Other” men</b>	-	-	-	-	-	-	-	-
<b>“Other” women</b>	5	0.2	3	0.1	4	0.1	4	0.1

### Study 2: Spring 2015

The participants, materials, and procedure in the Spring 2015 semester were identical to those described in the general method. A total of 643 participants completed the entire survey.

According to the binary gender and traditional sexual orientation question, the sample was 26% heterosexual men, 68% heterosexual women, 0.2% lesbian women, 0.9% gay men, 0.8% bisexual men, 2.2% bisexual women, 0.3% “other” men, and 0.5% “other” women.

**Results and Discussion**

**Hypothesis 1.**

Hypothesis 1 was less strongly supported than for Fall 2014, but most participants still seemed able and willing to answer the questions. The rate of missing responses was much higher (73 responses of 643, or 11.4%), while only 9 participants (1.4%) did not respond to the binary gender question. However, fewer participants (between 0.6% and 2.2% depending on the question) wrote in responses. Again, a small minority of participants selected the “Nobody” option and another option on each question. Most participants’ responses to the new sexual orientation question were congruent with their responses to the traditional sexual orientation question (see Hypothesis 2 for more detail). We tentatively concluded that most participants understood the questions, but we do not have an explanation for the dramatic increase in missing responses apart from variability in the other questions included in Mass Testing that semester.

**Hypothesis 2.**

Again, hypothesis 2 was supported. Using a liberal interpretation of their sexual identities, sexual attraction was misclassified for 6.7% of participants, and sexual behavior for 2.8%. With the conservative interpretation of their sexual identities, sexual attraction was misclassified for 0.6% of participants, and sexual behavior for 0.6%.

*Table 5. Inconsistent responses on sexual identity and sexual attraction/behavior – Spring 2015.*

	<b>Liberal definition</b>				<b>Conservative definition</b>			
	Attraction		Behavior		Attraction		Behavior	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>All participants</b>	43	6.7	18	2.8	4	0.6	4	0.6
<b>Heterosexual Men</b>	6	0.9	2	0.3	1	0.2	-	-

<b>Heterosexual Women</b>	25	3.9	12	1.9	3	0.5	2	0.3
<b>Gay Men</b>	1	0.2	2	0.3	1	0.2	2	0.3
<b>Lesbian Women</b>	-	-	-	-	-	-	-	-
<b>Bisexual Men</b>	4	0.6	2	0.3				
<b>Bisexual Women</b>	7	1.1	-	-				

### Hypotheses 3 and 4.

Hypotheses 3 and 4 were generally supported. Between 27.7% and 5.9% of participants had inconsistent responses on at least one of the four new sexual orientation questions, and there were participants with inconsistent answers across the four most plausible combinations of sexual and romantic attraction and behavior. As for Fall 2014, the numbers of people with inconsistent responses dropped dramatically between the liberal and conservative definition.

*Table 6. Inconsistent responses on new sexual orientation questions – Spring 2015.*

	<b>Liberal definition</b>		<b>Conservative definition</b>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
<b>Any inconsistent responses</b>	178	27.7	38	5.9
<b>Romantic and Sexual Attraction Inconsistent</b>	40	6.2	26	4.0
<b>Romantic and Sexual Behavior Inconsistent</b>	113	17.6	15	2.3
<b>Romantic Attraction and Behavior Inconsistent</b>	72	11.1	19	2.9
<b>Sexual Attraction and Behavior Inconsistent</b>	163	25.3	32	4.9

### Hypothesis 5.

Hypothesis 5 was supported; 3.7% of participants reported sexual/romantic attraction toward or sexual/romantic behavior with trans or non-binary people. Attraction (2.3-3.3%) was more common than behavior (0.5-0.6%). Proportions of attraction and behavior were roughly equivalent between trans men, trans women, and non-binary people, although no participants reported any romantic or sexual behavior with trans women. However, this might be due to a much smaller sample in Spring 2015.

*Table 7. Sexual/romantic attraction/behavior with non-cisgender people – Spring 2015 data.*

	<b>Any non-cis</b>		<b>Trans men</b>		<b>Trans women</b>		<b>Non-binary</b>	
	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>	<i>N</i>	<i>%</i>
<b>Any attraction/behavior</b>	24	3.7	17	2.6	14	2.2	13	2.0

<b>Romantic attraction</b>	15	2.3	10	1.6	10	1.6	11	1.7
<b>Sexual attraction</b>	21	3.3	13	2.0	11	1.7	11	1.7
<b>Romantic behavior</b>	4	0.6	3	0.5	-	-	1	0.2
<b>Sexual behavior</b>	3	0.5	2	0.3	-	-	1	0.2

### Hypothesis 6.

Hypothesis 6 was somewhat supported, although less obviously than in Fall 2014. Given the extremely small numbers of LGB people in the Spring 2014 sample (only 31 non-heterosexuals), several sexual identities were not represented among participants who reported sexual or romantic attraction or behavior with a trans or non-binary person. Again, this was more common among bisexual/pansexual people than among heterosexuals, and people from each identity category appeared to be attracted to/having relationships with trans people with a variety of gender identities and assigned sexes.

*Table 8. Any sexual/romantic attraction/behavior with non-cis people by sexual identity – Spring 2015.*

	<b>Any non-cis</b>		<b>Trans men</b>		<b>Trans women</b>		<b>Non-binary</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>Heterosexual men</b>	3	0.5	1	0.2	2	0.3	1	0.2
<b>Heterosexual women</b>	7	1.0	5	0.7	3	0.5	2	0.3
<b>Gay men</b>	-	-	-	-	-	-	-	-
<b>Lesbian women</b>	-	-	-	-	-	-	-	-
<b>Bisexual men</b>	4	0.6	3	0.5	-	-	2	0.3
<b>Bisexual women</b>	8	1.2	7	1.0	7	1.0	7	1.0
<b>“Other” men</b>	-	-	-	-	-	-	-	-
<b>“Other” women</b>	2	0.3	1	0.2	2	0.3	-	-

### Study 3: Fall 2015

The participants, materials, and procedure in the Fall 2015 semester were identical to those described in the general method. A total of 1818 participants completed the entire survey. According to the binary gender and traditional sexual orientation question, the sample was 38% heterosexual men, 55% heterosexual women, 0.3% lesbian women, 1.5% gay men, 1.2% bisexual men, 2.8% bisexual women, 0.1% “other” men, and 0.6% “other” women.

## Results and Discussion

### Hypothesis 1.

Hypothesis 1 was supported; most participants still seemed able and willing to answer the questions. There were 25 missing responses (1.4%), which is similar to Fall 2014, much lower than for Spring 2015, and roughly equivalent to the number who did not respond to the binary gender question (23, or 1.3%). Between 2.8% and 3.6% of participants wrote in responses, similar to the proportion for the previous two semesters, and again, a small minority of participants selected the “Nobody” option and another option on each question. Most participants’ responses to the new sexual orientation question were congruent with their responses to the traditional sexual orientation question (see Hypothesis 2 for more detail). We concluded that most participants understood the questions.

### Hypothesis 2.

Hypothesis 2 was supported. Using a liberal interpretation of their sexual identities, sexual attraction was misclassified for 9.6% of participants, and sexual behavior for 6.5%. With the conservative interpretation of their sexual identities, sexual attraction was misclassified for 0.6% of participants, and sexual behavior for 0.4%.

*Table 9. Inconsistent responses on sexual identity and sexual attraction/behavior – Fall 2015.*

	Liberal definition				Conservative definition			
	Attraction		Behavior		Attraction		Behavior	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>All participants</b>	174	9.6	118	6.5	10	0.6	8	0.4
<b>Heterosexual Men</b>	53	2.9	39	2.1	3	0.2	5	0.3
<b>Heterosexual Women</b>	76	4.2	61	3.4	6	0.3	2	0.1
<b>Gay Men</b>	5	0.3	6	0.3	1	0.1	1	0.1
<b>Lesbian Women</b>	1	0.1	2	0.1	-	-	-	-
<b>Bisexual Men</b>	9	0.5	6	0.3				
<b>Bisexual Women</b>	30	1.7	4	0.2				

### Hypotheses 3 and 4.

Hypotheses 3 and 4 were generally supported. Between 28.4% and 9.9% of participants had inconsistent responses on at least one of the four new sexual orientation questions, and there were participants with inconsistent answers across the four most plausible combinations of sexual and romantic attraction and behavior. As for the previous two semesters, the numbers of people with inconsistent responses dropped dramatically between the liberal and conservative definition.

*Table 10. Inconsistent responses on new sexual orientation questions – Fall 2015 data.*

	<b>Liberal definition</b>		<b>Conservative definition</b>	
	<i>N</i>	%	<i>N</i>	%
<b>Any inconsistent responses</b>	516	28.4	181	9.9
<b>Romantic and Sexual Attraction Inconsistent</b>	157	8.6	114	6.3
<b>Romantic and Sexual Behavior Inconsistent</b>	331	18.2	91	5.0
<b>Romantic Attraction and Behavior Inconsistent</b>	234	12.8	98	5.4
<b>Sexual Attraction and Behavior Inconsistent</b>	438	24.0	136	7.5

### **Hypothesis 5.**

Hypothesis 5 was supported; 6.1% of participants reported sexual/romantic attraction toward or sexual/romantic behavior with trans or non-binary people. Attraction (3.9-4.3%) was more common than behavior (1.6-1.8%). As before, proportions of attraction and behavior were roughly equivalent between trans men, trans women, and non-binary people.

*Table 11. Sexual/romantic attraction/behavior with non-cisgender people – Fall 2015.*

	<b>Any non-cis</b>		<b>Trans men</b>		<b>Trans women</b>		<b>Non-binary</b>	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>Any attraction/behavior</b>	111	6.1	76	4.2	58	3.2	58	3.2
<b>Romantic attraction</b>	70	3.9	54	3.0	34	1.9	42	2.3
<b>Sexual attraction</b>	78	4.3	50	2.8	36	2.0	39	2.1
<b>Romantic behavior</b>	33	1.8	10	0.6	10	0.6	17	0.9
<b>Sexual behavior</b>	29	1.6	9	0.5	4	0.2	17	0.9

### **Hypothesis 6.**

Hypothesis 6 was supported. Participants from almost all sexual identities reported sexual or romantic attraction or behavior with a trans or non-binary person. Again, this was more

common among bisexual/pansexual people than among heterosexuals, and people from each identity category appeared to be attracted to/having relationships with trans people with a variety of gender identities and assigned sexes.

*Table 12. Any sexual/romantic attraction/behavior with non-cis people by sexual identity - Fall 2015.*

	Any non-cis		Trans men		Trans women		Non-binary	
	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	%
<b>Heterosexual men</b>	31	1.7	11	0.6	16	0.9	11	0.6
<b>Heterosexual women</b>	24	1.3	18	1.0	9	0.5	6	0.3
<b>Gay men</b>	1	0.1	4	0.2	1	0.1	1	0.1
<b>Lesbian women</b>	5	0.2	-	-	-	-	2	0.1
<b>Bisexual men</b>	10	0.6	8	0.4	5	0.3	5	0.3
<b>Bisexual women</b>	29	1.6	27	1.5	20	1.1	25	1.4
<b>“Other” men</b>	1	0.1	1	0.1	1	0.1	-	-
<b>“Other” women</b>	7	0.3	5	0.3	4	0.2	6	0.3

## General Discussion

### Summary

In three large samples totaling over 5000 participants, we found consistent support for all of our hypotheses. This provides preliminary evidence that sexual orientation questions asking about sexual and romantic attraction and behavior toward a variety of genders are practical and provide information that is missed and/or muddled by current identity-only measures. We found that most participants were able to answer the questions, as demonstrated by small numbers of write-in responses and overall similarity between responses to the traditional sexual orientation question and the new ones. Although one sample had an 11% missing response rate, the other two samples had very small rates of missing data, suggesting that participants are willing to provide this information.

Across samples, the constructs of attraction and behavior were consistently distinct from identity. While most participants’ identities were predictably related to the genders to whom they were attracted and with whom they engaged in relationships, a significant minority (3-6%)

showed inconsistency in this regard even when using a very conservative standard. For perspective, the number of people whose sexual attractions or behaviors included a cisgender person of a gender normally considered “excluded” by their identity (e.g. heterosexual men with cisgender men) was equivalent to or larger than the number of lesbian women (around 0.5%). As expected from previous research (e.g. Rosario & Schrimshaw, 2014), attraction and behavior were distinct from each other. Like previous asexuality and bisexuality researchers (e.g. Weinberg et al., 1994; Bogaert, 2013), we also found evidence that the sexual and romantic dimensions of these constructs differed.

Of central importance, 3-6 percent of participants indicated that they experienced sexual or romantic attraction or behavior with trans people. These participants had a variety of sexual identities, and experienced attraction and/or behavior with trans men, trans women, and non-binary people and roughly equivalent rates. This suggests that measuring sexual attraction to and behavior with transgender people and those with non-binary identities is immediately feasible, and that these experiences are relatively common (similar to holding an LGB identity).

### **Current Recommendations**

For anyone looking to immediately implement an inclusive sexual orientation measure, we wanted to offer some concrete recommendations based on our findings from this research and our other studies on transgender-inclusive gender demographic questions.

First, we recommend including a sexual identity question. If resources for coding responses are not a concern, this could be a free-text question. Alternatively, a multiple-choice question would also be feasible. We recommend the following measure, adapted from one used by the National Transgender Discrimination Survey (Grant et al., 2011), for further testing:

*What term(s) do you usually use to describe your sexual identity? Please select all that apply.*

- *Asexual*
- *Aromantic*



- *Bisexual*
- *Gay*
- *Heterosexual*
- *Lesbian*
- *Pansexual*
- *Queer*
- *Additional term(s) not listed [free text box]*
- *Do not know*
- *Choose not to answer*

Next, we recommend adapting our attraction and behavior questions to include definitions for “cisgender,” “transgender,” and “non-binary.” If desired, researchers could also specify a time period and/or ask participants how stable/fluid they perceive their identity to be (as sexualities can change over time; e.g. Diamond, 2008). We also suggest making the “Nobody” option exclusive (such that participants cannot select it and other options), changing the wording of the “Other” option to “Another gender not listed” (as the wording is less othering; cf. Balarajan et al., 2011), changing the “Genderqueer” option to “Non-binary people (e.g. genderqueer people),” and including “Do not know” and “Choose not to answer” options. The revised set of measures might look like this (only the example for sexual attraction is shown):

*Definition of terms: **Non-binary** refers to a person whose gender identity is something other than man or woman. **Cisgender** refers to a person whose gender identity is the same as the sex assigned to them at birth, while **transgender** refers to a person whose gender identity is different from the sex assigned to them at birth. For example, a **cisgender** man is a person who was assigned male at birth and who identifies as a man, and a **transgender** man is a person who was assigned female at birth and who identifies as a man.*

*In the past year, what genders of people have you been sexually attracted to? Select all that apply.*

- *Cisgender men*
- *Transgender men*
- *Cisgender women*
- *Transgender women*
- *Non-binary people (e.g. genderqueer people)*
- *Another gender not listed [free text]*
- *I have not been sexually attracted to anyone in the past year.*
- *I do not know.*
- *I choose not to answer.*

## **Limitations**

The research presented here has several limitations. It relies on a convenience sample of undergraduate students, who presumably have a relatively narrow range of ages and sexual experiences, and the results may not generalize well to other populations. We also have very

small numbers of participants in some identity categories, and a large number of missing responses, so these results should be treated as exploratory and suggestive. This research should be replicated with more diverse samples, and if possible, future researchers should ask about attraction and behavior in different time periods and/or follow participants longitudinally to gain a better understanding of the patterns and extent of sexual fluidity. As we did not explicitly define the terms we used, it is possible that participants interpreted them differently than we did. Defining terms is one way to reduce this concern (as described in the previous section), but cognitive testing and/or focus groups with a diverse group of participants would also give us more confidence in our results. More importantly, qualitative research might help us identify yet more unmet needs and new ways of conceptualizing sexuality.

We are also well aware that the measures presented here barely scratch the surface of what is needed to make sexuality research truly inclusive, and understand that incorporating currently marginalized sexualities into mainstream theory and practice will take decades. For example, while van Anders (2015) provides an extensive summary of dozens of currently recognized problems with our conceptualization of sexualities and one possible integrated approach, it remains too unwieldy for use in survey contexts as yet. We have chosen to take an incremental approach to creating more inclusive sexual orientation questions in the hopes of fostering more rapid (if smaller) improvements to the status quo. Thus, we fully expect this measure to be supplanted by an improved model in the near future.

### **Future Research Directions**

One particularly promising direction for future research might include more extensive use of branching survey logic. For example, participants could be asked whether they experience sexual attraction at all, and whether they experience attraction based on gender, with only those

who agree proceeding to a question where they indicate which genders they are attracted to. Similar designs could be used to ask what aspects of gender participants find attractive (identities, bodies, etc.) Branching logic would enable clarification of otherwise ambiguous responses such as “I don’t know,” for instance by asking whether the participant does not understand some aspect of the question, or understands the question but does not know how to translate their own experience into the categories provided by the question, potentially increasing the amount of useable data (Cain, 2012). With more complex designs, participants who indicate particularly interesting patterns of responses on the attraction and behavior questions (for example, those with inconsistent responses) could be asked follow-up questions to clarify the meaning of their answers. Future researchers might also want to use branching logic to ask more detailed questions about the nature of participants’ attractions to and sexual behaviors with transgender people, as these questions could be displayed only when applicable. For example, it would be interesting to know whether participants’ transgender partners already identified as such at the time of the attraction/relationship/sexual encounter, and whether/how the participants changed their sexual identities as a result of these interactions.

Alongside this work on question design, researchers who are interested in sexuality per se should begin to confront the limitations and imprecision of standard measures and conceptualizations of sexual orientation, and begin an empirical investigation of the relationships between the constellation of dimensions that influence gender and desire.

First, researchers must specify the relevant qualities of the agent who does the desiring and the target who is desired. If we believe that a person’s gender influences their own desires, and determines who will desire them, which precise aspect(s) of gender are relevant? When people conceive of themselves as gendered beings, their private identities, social identities,

physical presentations, and physical bodies each seem to exert their own influence (Galupo et al., 2014; Tate et al., 2014), and this list is likely far from exhaustive. What roles might each of these components of gender play in shaping the ways people desire? As yet, we have almost no empirical knowledge on the subject. We might next wonder where the relevant desire is located. While many people conceive of their sexual identities as originating within the self (Galupo et al., 2014), the construct of sexual orientation is inherently relational, emerging from the relationship between one person's identity and another's. Which components of sexual identity can properly be described as self-defined, and which are socially or relationally defined? How and when do these definitions emerge? Are they stable or fluid over time? These questions also have yet to be answered.

A growing number of researchers are even returning to the question of what we mean by desire in the first place. Even comparatively early sexual orientation researchers (e.g. Klein, 1985) drew distinctions between sexual behavior, sexual attraction, sexual fantasy, emotional preference, and social preference, but these different constructs have largely been excluded from the public imagination in favor of an increased focus on a single, unifying sexual orientation. Asexuality activists have called for researchers to renew our attention to these distinctions, suggesting that people may have different sexual and romantic orientations and that these distinctions are significant and worthy of study (e.g. Bogaert, 2013). Although researchers are beginning to consider these questions, progress is slow.

Finally, queer theorists and scholars of bisexuality (and other non-monosexual identities) have been willing to question not just the ways in which sexual orientation conceptualizes gender, but the idea that gender is the central locus of desire in the first place. If researchers step back from the assumption that desire is always and exclusively oriented by gender, what other

dimensions of desire might we uncover? Consider the identities being constructed within kink and BDSM communities around power (Dominant/submissive roles) and sensation (top/bottom, sadist/masochist, giving/receiving sensation; e.g. Faccio et al., 2014; Hebert & Weaver, 2015). Might these dimensions be conceptualized in similar ways to gendered desires? These other potential dimensions of desire have received very little in the way of scholarly attention, but it's possible that by considering multiple aspects of sexual desire, attraction, and behavior we can gain a better understanding of what motivates people to behave the way they do, and perhaps eventually predict behavior with some accuracy.

### **Conclusion**

Despite decades of research from multiple disciplines which complicates binary views of gender, the construct of sexual orientation and measures for assessing it continue to rely on the inaccurate assumption that there exist two mutually exclusive, exhaustive, readily definable gender categories. As a result of this systematic failure to incorporate new insights into practice, researchers systematically misclassify and outright erase the experiences of transgender people, people with non-binary gender identities, and their partners. This exclusion results in lower-quality health care for members of some of the most marginalized communities (e.g. trans women) and may exacerbate their feelings of stigma. Furthermore, research and practice that relies on sexual orientation as a proxy for other, more specific information does a disservice to marginalized communities and diminishes the accuracy and generalizability of data, hampering progress toward a scientific understanding of the relationship between gender and sexuality. To address these problems, clinicians and other practitioners who need quick data should replace sexual orientation questions with more targeted questions that more appropriately operationalize the construct of interest.

Sexual orientation questions which include transgender, non-binary, and asexual people are greatly needed. We created a set of inclusive questions about sexual and romantic attraction and behavior and tested them with three large undergraduate samples. Participants were able to answer the questions, and their responses matched established theory suggesting that sexual identity, attraction, and behavior are separate constructs. A substantial minority also indicated that they experienced attraction to or behavior with transgender or non-binary people. We offer a revised version of our measure for immediate use by future researchers, and detail several promising new directions for the design of inclusive sexual orientation questions (particularly through the use of branching survey logic).

## References

- American Psychological Association (APA). (2010). *Publication manual of the American Psychological Association*. Washington, D.C.: American Psychological Association.
- American Psychological Association (APA). Task Force on Gender Identity and Gender Variance. (2009). *Report of the Task Force on Gender Identity and Gender Variance*. Washington, DC: American Psychological Association. Retrieved from [www.apa.org/pi/lgbc/transgender/2008TaskForceReport.html](http://www.apa.org/pi/lgbc/transgender/2008TaskForceReport.html)
- American Public Health Association. (1999). APHA policy statement 9819: the need for public health research on gender identity and sexual orientation. *American Journal of Public Health*, 89. 444–445.
- Ansara, Y. G., & Hegarty, P. (2012). Cisgenderism in psychology: pathologising and misgendering children from 1999 to 2008. *Psychology & Sexuality*, 3(2), 137-160. doi:10.1080/19419899.2011.576696
- Ard, K.L., & Makadon, H.J. (2012). *Improving the health care of lesbian, gay, bisexual and transgender (LGBT) people: Understanding and eliminating health disparities*. Boston, MA: The Fenway Institute. Retrieved from [http://www.lgbthealtheducation.org/wp-content/uploads/12-054\\_LGBTHealtharticle\\_v3\\_07-09-12.pdf](http://www.lgbthealtheducation.org/wp-content/uploads/12-054_LGBTHealtharticle_v3_07-09-12.pdf)
- Balarajan, M., Gray, M. & Mitchell, M. (2011). *Monitoring equality: Developing a gender identity question*. (Equality and Human Rights Commission Research Report 75). Retrieved from <http://www.equalityhumanrights.com/publications/our-research/research-reports/research-reports-71-80/>
- Bancroft, J. (2009). *Human Sexuality and its Problems* (3rd edition). London: Churchill Livingstone/Elsevier.



- Bargh, J. A., Lombardi, W.J., & Higgins, E.T. (1988). Automaticity of Person x Situation effects on impression formation: It's just a matter of time. *Journal of Personality and Social Psychology*, 55(4), 599–605.
- Baron, D. (2018, September 4). A brief history of singular 'they.' *Oxford English Dictionary Blog*. Retrieved from <https://public.oed.com/blog/a-brief-history-of-singular-they/>.
- Bauer, G. R., Hammond, R., Travers, R., Kaay, M., Hohenadel, K. M., & Boyce, M. (2009). "I Don't Think This Is Theoretical; This Is Our Lives": How Erasure Impacts Health Care for Transgender People. *Journal of the Association of Nurses in AIDS Care*, 20(5), 348–361. doi:10.1016/j.jana.2009.07.004
- Bauerband, L. A., & Galupo, M. P. (2014). The Gender Identity Reflection and Rumination Scale: Development and Psychometric Evaluation. *Journal of Counseling & Development*, 92(2), 219–231. <https://doi.org/10.1002/j.1556-6676.2014.00151.x>
- Beh, H., & Diamond, M. (2005). Ethical concerns related to treating gender nonconformity in childhood and adolescence: Lessons from the Family Court of Australia. *Health Matrix: Journal of Law-Medicine*, 15, 239-283
- Bettcher, T. M. (2007). Evil Deceivers and Make-Believers : On Transphobic Violence and the Politics of Illusion. *Hypatia*, 22(3), 43–66. doi:10.1111/j.1527-2001.2007.tb01090.x
- Blackless, M., Charuvastra, A., Derrtyck, A., Fausto-Sterling, A., Lauzanne, K., & Lee, E. (2000). How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology*, 12, 151–166.
- Blashill, A. J., & Powlishta, K. K. (2009). Gay stereotypes: The use of sexual orientation as a cue for gender-related attitudes. *Sex Roles*, 61, 783–793.

- Bogaert, A. (2013). The Demography of Asexuality. In A.K. Baumle (Ed.), *International handbook on the demography of sexuality*. (pp. 21-43). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-5512-3.
- Bornstein, K. (1994). *Gender Outlaw: On Men, Women, and The Rest of Us*. New York, NY: Vintage.
- Brehm, J.W. (1966). *A theory of psychological reactance*. New York: Academic Press.
- Broussard, K. A., Warner, R. H., & Pope, A. R. D. (2017). Too Many Boxes, or Not Enough? Preferences for How We Ask About Gender in Cisgender, LGB, and Gender-Diverse Samples. *Sex Roles*, 1–19. <https://doi.org/10.1007/s11199-017-0823-2>
- Buss, D. M. (1995). Psychological sex differences: Origins through sexual selection. *American Psychologist*, 50, 164-168.
- Bussey, K. (2011). Gender identity development. In S. J. Schwartz, K. Lyuckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (p. 603–628). New York, NY: Springer. [http://dx.doi.org/10.1007/978-1-4419-7988-9\\_26](http://dx.doi.org/10.1007/978-1-4419-7988-9_26)
- Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, 106(4), 676–713. doi:10.1016/j.sbspro.2011.03.261
- Butler, J. (1990). *Gender trouble: feminism and the subversion of identity*. New York: Routledge.
- Butler, J. (2006). *Gender trouble: Feminism and the subversion of identity*. New York: Routledge.
- Cahill, S., & Makadon, H. (2014). *Sexual Orientation and Gender Identity Data Collection in Clinical Settings and in Electronic Health Records* : (Vol. 1). doi:10.1089/lgbt.2013.0001

- Cain, V.S. (2012). *Improving measurement of sexual orientation in US health surveys*. Paper presented at the European Conference on Quality in Official Statistics. Retrieved from [http://www.q2012.gr/articlefiles/sessions/29.1\\_%20Cain\\_Sexual%20Orientation%20CAIN.pdf](http://www.q2012.gr/articlefiles/sessions/29.1_%20Cain_Sexual%20Orientation%20CAIN.pdf)
- Califia, P. (1997). *Sex changes: the politics of transgenderism*. San Francisco, CA: Cleis Press.
- Cheryan, S., & Monin, B. (2005). “Where are you really from?": Asian Americans and identity denial. *Journal of Personality and Social Psychology*, *89*(5), 717–730.  
<https://doi.org/10.1037/0022-3514.89.5.717>
- Ching, B. H. H., & Xu, J. T. (2017). The Effects of Gender Neuroessentialism on Transprejudice: An Experimental Study. *Sex Roles*, 1–14. <https://doi.org/10.1007/s11199-017-0786-3>
- Cohen-Kettenis, P. T., & Pfäfflin, F. (2010). The DSM diagnostic criteria for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior*, *39*, 499–513.  
[doi:10.1007/s10508-009-9562-y](https://doi.org/10.1007/s10508-009-9562-y)
- Coleman, E., Bockting, W., Botzer, M., Cohen-Kettenis, P., DeCuypere, G., Feldman, J.,...Zucker, K. (2012). Standards of care for the health of transsexual, transgender, and gender non-conforming people, Version 7. *International Journal of Transgenderism*, *13*(4), 165-232, <https://doi.org/10.1080/15532739.2011.700873>
- Coleman, J. M., & Hong, Y.-Y. (2008). Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity*, *7*(1), 34–53.  
<http://doi.org/10.1080/15298860600980185>

- Conron, K. J., Scott, G., Stowell, G. S., & Landers, S. J. (2012). Transgender health in massachusetts: Results from a household probability sample of adults. *American Journal of Public Health, 102*(1), 118–122. doi:10.2105/AJPH.2011.300315
- Conron, K., Lombardi, E., & Reisner, S. (2014). Identifying transgender and other gender minority respondents on population-based surveys: Approaches. In J.L. Herman (Ed.), *Best practices for asking questions to identify transgender and other gender minority respondents on population-based surveys*. (p. 9-18). Los Angeles, CA: The Williams Institute.
- Consortium on the Management of Disorders of Sex Development. (2006). *Clinical Guidelines for the Management of Disorders of Sex Development in Childhood*. (1<sup>st</sup> ed.). Rohnert Park, CA: Intersex Society of North America.
- Cotton, A. (2013, November 8). Colorado College defends use of "queer" on application. *The Denver Post*. Retrieved from [www.denverpost.com/news/ci\\_24478848/colorado-college-defends-use-queer-application](http://www.denverpost.com/news/ci_24478848/colorado-college-defends-use-queer-application)
- Dank, M., Lachman, P., Zweig, J. M., & Yahner, J. (2014). Dating violence experiences of lesbian, gay, bisexual, and transgender youth. *Journal of Youth and Adolescence, 43*(5), 846–857. <http://doi.org/10.1007/s10964-013-9975-8>
- de Lemus, S., Bukowski, M., Spears, R., & Telga, M. (2015). Reactance to (or acceptance of) stereotypes: Implicit and explicit responses to group identity threat. *Zeitschrift Fur Psychologie / Journal of Psychology, 223*(4), 236–246. <https://doi.org/10.1027/2151-2604/a000225>

- Department of Health and Human Services. *2020 Topics and Objectives: Lesbian, Gay, Bisexual, and Transgender Health*. (2013, August 28). Retrieved from <http://www.healthypeople.gov/2020/topicsobjectives2020/objectiveslist.aspx?topicId=25>.
- Deutsch, M. B., Green, J., Keatley, J., Mayer, G., Hastings, J., & Hall, A. M. (2013). Electronic medical records and the transgender patient: recommendations from the World Professional Association for Transgender Health EMR Working Group. *Journal of the American Medical Informatics Association : JAMIA*, *20*, 700–3. doi:10.1136/amiajnl-2012-001472
- Devor, A. (2004). Witnessing and mirroring: A fourteen stage model of transsexual identity formation. *Journal of Gay & Lesbian Mental Health*, *8*(1), 41–67. doi:10.1080/19359705.2004.9962366
- deVries, A.L.C., Kreukles, B.P.C., Steensma, T.D., & McGuire, J.K. (2014). Gender identity development: A biopsychosocial perspective. In Kreukles, B.P.C., Steensma, T.D., and deVries, A.L.C. (Eds.), *Gender Dysphoria and Disorders of Sex Development: Progress in Care and Knowledge*.
- Diamond, L. (2003). What Does Sexual Orientation Orient? A Biobehavioral Model Distinguishing Romantic Love and Sexual Desire. *Psychological Review*, *110*(1), 173–192.
- Diamond, L. M. (2008). Female bisexuality from adolescence to adulthood: results from a 10-year longitudinal study. *Developmental Psychology*, *44*(1), 5–14. <http://doi.org/10.1037/0012-1649.44.1.5>
- Diamond, M. (2000). IV. Sex and Gender: Same or Different? *Feminism & Psychology*, *10*(1), 46–54. doi:10.1177/0959353500010001007

- Diamond, L.M., Pardo, S.T., & Butterworth, M.R. (2011). Transgender identity and experience. In Schwartz, S.J., Luyckx, K., & Vignoles, V.L. (Eds.), *Handbook of Identity Theory and Research* (pp. 629-647). New York, NY: Springer. doi: 10.1007/978-1-4419-7988-9\_26
- Dreger, A. D. (1998). 'Ambiguous Sex': Or Ambivalent Medicine? Ethical Issues in the Treatment of Intersexuality. *The Hastings Center Report*, (3)24.
- Drescher, J. (2010). Queer diagnoses: Parallels and contrasts in the history of homosexuality, gender variance, and the diagnostic and statistical manual. *Archives of Sexual Behavior*, 39, 427–460. [http://dx.doi.org/ 10.1007/s10508-009-9531-5](http://dx.doi.org/10.1007/s10508-009-9531-5)
- Durso, L.E. & Gates, G.J. (2013). Best practices: Collecting and analyzing data on sexual minorities. In A.K. Baumle (Ed.), *International handbook on the demography of sexuality*. (pp. 21-43). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-5512-3.
- Dweck, C. S., Chiu, C., & Hong, Y. (1995). Implicit Theories and Their Role in Judgments and Reactions: A Word From Two Perspectives. *Psychological Inquiry*, 6(4), 267–285. [http://doi.org/10.1207/s15327965pli0604\\_1](http://doi.org/10.1207/s15327965pli0604_1)
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.
- Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, 37, 451–463.
- Ehrensaft, D. (2012). From gender identity disorder to gender identity creativity: True gender self child therapy. *Journal of Homosexuality*, 59, 337–356.
- Equal Employment Opportunity Commission (EEOC). (2018, October 13). Pre-Employment Inquiries and Gender. Retrieved from [https://www.eeoc.gov/laws/practices/inquiries\\_gender.cfm](https://www.eeoc.gov/laws/practices/inquiries_gender.cfm)

- Faccio, E., Casini, C., & Cipolletta, S. (2014). Forbidden games: the construction of sexuality and sexual pleasure by BDSM “players.” *Culture, Health & Sexuality*, 16(7), 752–764.
- Factor, R., & Rothblum, E. (2008). Exploring gender identity and community among three groups of transgender individuals in the United States: MTFs, FTMs, and genderqueers. *Health Sociology Review*, 17, 235–253. <http://dx.doi.org/10.5172/hesr.451.17.3.235>
- Fausto-Sterling, A. (2000). *Sexing the body: gender politics and the construction of sexuality*. New York, NY: Basic Books.
- Fausto-Sterling, A. (2012). The dynamic development of gender variability. *Journal of Homosexuality*, 59, 398–421. <http://dx.doi.org/10.1080/00918369.2012.653310>
- Fenway Institute. (2013). *Asking Patients Questions about Sexual Orientation and Gender Identity in Clinical Settings: A Study in Four Health Centers*. Report of the Fenway Institute and the Center for American Progress. Retrieved from [http://thefenwayinstitute.org/wp-content/uploads/COM228\\_SOGI\\_CHARN\\_WhitePaper.pdf](http://thefenwayinstitute.org/wp-content/uploads/COM228_SOGI_CHARN_WhitePaper.pdf).
- Food and Drug Administration (FDA). (2015). *Blood Donations from Men Who Have Sex with Other Men: Questions and Answers*. Retrieved from [www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/QuestionsaboutBlood/ucm108186.htm](http://www.fda.gov/BiologicsBloodVaccines/BloodBloodProducts/QuestionsaboutBlood/ucm108186.htm)
- Freud, S. (1962). *Three contributions to the theory of sex* (J. Strachey, Trans.). New York: Nervous and Mental Disease. (Original work published 1905).
- Galupo, M. P., Mitchell, R. C., Gryniewicz, A. L., Davis, K. S., & Gryniewicz, A. L. (2014). Conceptualization of Sexual Orientation Identity Among Sexual Minorities: Patterns Across

Sexual and Gender Identity. *Journal of Bisexuality*, 14 (3-4), 433-456,

doi:10.1080/15299716.2014.933466

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice-Hall.

Garofalo, R., Deleon, J., Osmer, E., Doll, M., & Harper, G. W. (2006). Overlooked, misunderstood and at-risk: Exploring the lives and HIV risk of ethnic minority male-to-female transgender youth. *Journal of Adolescent Health*, 38(3), 230–236.

<http://doi.org/10.1016/j.jadohealth.2005.03.023>

GenIUSS Group, The. (2014). *Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys*. J.L. Herman (Ed.). Los Angeles, CA: The Williams Institute.

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70, 491-512.

Graham-Bailey, M., Richardson Cheeks, B. L., Blankenship, B. T., Stewart, A. J., & Chavous, T. M. (2018, December 10). Examining College Students' Multiple Social Identities of Gender, Race, and Socioeconomic Status: Implications for Intergroup and Social Justice Attitudes. *Journal of Diversity in Higher Education*. Advance online publication. <http://dx.doi.org/10.1037/dhe0000098>

Grant, J.M., Mottet, L.A., Tanis, J., Harrison, J., Herman, J.L., & Keisling, M. (2011). *Injustice at Every Turn: A Report of the National Transgender Discrimination Survey*. Washington: National Center for Transgender Equality and National Gay and Lesbian Task Force.

Grant, J.M., Mottet, L.A., Tanis, J., Harrison, J., Herman, J.L., & Keisling, M. (2011). *Injustice at Every Turn: A Report of the National Transgender Discrimination Survey*.



Washington: National Center for Transgender Equality and National Gay and Lesbian Task Force.

Guo, J. (2016, January 8). Sorry, grammar nerds. The singular ‘they’ has been declared Word of the Year. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/news/wonk/wp/2016/01/08/donald-trump-may-win-this-years-word-of-the-year/>

Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *The British Journal of Social Psychology / the British Psychological Society*, 39 (Pt 1), 113–127. <https://doi.org/10.1348/014466600164363>

Hauptert, M. (2015). *Is the gender binary a sacred value?: Moral dimensions of transphobia*. Unpublished manuscript, Indiana University, Bloomington, IN.

Hauptert, M.L., Murphy, M.C. & Smith, E.R. (2014). *How context and wording of gender questions affect perceptions of researchers* (unpublished manuscript). Indiana University: Bloomington, Bloomington, IN, USA.

Hébert, A., & Weaver, A. (2015). An Examination of Personality Characteristics Associated with BDSM Orientations. *The Canadian Journal of Human Sexuality*, 24(1), 49–62. doi:10.3138/cjhs.2467

Hegarty, P., & Pratto, F. (2010). Interpreting and communicating the results of gender-related research. In: J.C. Chrisler and D.R. McCreary (Eds.). *Handbook of gender research in psychology: Vol. 1. Gender research in general and experimental psychology*. New York: Springer, 191-211.

Herek, G.M., Kimmel, D.C., Amaro, H., & Melton, G.B. (1991). Avoiding heterosexist bias in psychological research. *American Psychologist*, 46(9), 957-963.

- Hill, D. B., & Willoughby, B. L. B. (2005). The Development and Validation of the Genderism and Transphobia Scale. *Sex Roles, 53*(7–8), 531–544. <http://doi.org/10.1007/s11199-005-7140-x>
- Hines, S., & Sanger, T. (2010). *Transgender Identities: Toward a Social Analysis of Gender Diversity*. New York, NY: Routledge.
- Hirschfeld, M. (1923). Die intersexuelle Konstitution. *Jahrbuch fur Sexuelle Zwischenstufen, 23*, 3–27.
- Horn, S. S. (2007). Adolescents' acceptance of same-sex peers based on sexual orientation and gender expression. *Journal of Youth & Adolescence, 36*, 363–371.
- Hubbard, R. (1996). Gender and Genitals: Constructs of Sex and Gender. *Social Text, 46*, 157–165.
- Institute of Medicine. (2011). *The Health of Lesbian, Gay, Bisexual, and Transgender People: Building a Foundation for Better Understanding*. Washington, DC: The National Academies Press. Retrieved from <http://www.iom.edu/Reports/2011/The-Health-of-Lesbian-Gay-Bisexual-and-Transgender-People.aspx>.
- James, S. E., Herman, J. L., Rankin, S., Keisling, M., Mottet, L., & Anafi, M. (2016). The Report of the 2015 U.S. Transgender Survey. Washington, DC: National Center for Transgender Equality.
- Joel, D., Tarrasch, R., Berman, Z., Mukamel, M., & Ziv, E. (2014). Queering gender: Studying gender identity in “normative” individuals. *Psychology & Sexuality, 5*(4), 291-321. doi:10.1080/19419899.2013.830640
- John, O.P., & Benet-Martinez, V. (2000). Measurement: Reliability, construct validation, and scale construction. In Reis, H.T., & Judd, C.M. (Eds.), *Handbook of research methods in*

- social and personality psychology*, (2<sup>nd</sup> ed., pp. 473-503). New York, NY: Cambridge University Press.
- Johnson, T. W., & Wassersug, R. J. (2010). Gender identity disorder outside the binary: When gender identity disorder-not otherwise specified is not good enough. *Archives of Sexual Behavior*, *39*, 597–598. doi:10.1007/s10508-010-9608-1
- Kessler, S. J. (1990). The Medical Construction of Gender: Case Management of Intersexed Infants. *Signs: Journal of Women in Culture and Society*, *16*(1), 3. doi:10.1086/494643
- Kessler, S. J., & McKenna, W. (1978). *Gender: An ethnomethodological approach*. Chicago, IL: University of Chicago Press.
- Kinsey, A.C., Pomeroy, W.B., & Martin, C.E. (1948). *Sexual Behavior in the Human Male*. W.B. Saunders, Philadelphia, PA.
- Klein, F., Sepekoff, B., & Wolf, T.J. (1985). Sexual orientation: a multi-variable dynamic process. *Journal of Homosexuality*, *11*, 35-49.
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82-173). Stanford, CA: Stanford University Press.
- Kozee, H. B., Tylka, T. L., & Bauerband, L. a. (2012). Measuring Transgender Individuals' Comfort With Gender Identity and Appearance: Development and Validation of the Transgender Congruence Scale. *Psychology of Women Quarterly*, *36*(2), 179–196. <https://doi.org/10.1177/0361684312442161>
- Kraus, M. W., Chen, S., & Keltner, D. (2011). The power to be me: Power elevates self-concept consistency and authenticity. *Journal of Experimental Social Psychology*, *47*(5), 974–980. <https://doi.org/10.1016/j.jesp.2011.03.017>

- Kuper, L. E., Nussbaum, R., & Mustanski, B. (2012). Exploring the Diversity of Gender and Sexual Orientation Identities in an Online Sample of Transgender Individuals. *Journal of Sex Research, 49*(2-3), 244–254. doi:10.1080/00224499.2011.596954
- Langone, A. (2018). #MeToo and Time's Up Founders Explain the Difference Between the 2 Movements — And How They're Alike. *TIME*. Retrieved from <http://time.com/5189945/whats-the-difference-between-the-metoo-and-times-up-movements/>
- Lawrence, A. A. (2010). Sexual orientation versus age of onset as bases for typologies (subtypes) for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior, 39*, 514–545. doi:10.1007/s10508-009-9594-3
- Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., ... Spears, R. (2008). Group-level self-definition and self-investment: a hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology, 95*(1), 144–165. <https://doi.org/10.1037/0022-3514.95.1.144>
- Lombardi, E. L., Banik, S., Mitchell, K., & Zuber, J. (2013). *Examination of gender identity and expression measures within a Midwest sample*. Retrieved from [http://www.researchgate.net/publication/258112182\\_Examination\\_of\\_Gender\\_Identity\\_and\\_Expression\\_Final\\_Report\\_for\\_Williams\\_Institute](http://www.researchgate.net/publication/258112182_Examination_of_Gender_Identity_and_Expression_Final_Report_for_Williams_Institute).
- Luhtanen, R., & Crocker, J. (1992). A Collective Self-Esteem Scale: Self-Evaluation of One's Social Identity. *Personality and Social Psychology Bulletin, 18*(3), 302–318. <https://doi.org/10.1177/0146167292183006>
- Martin, C.L. & Halverson, C.F. (1995). A schematic processing model of sex typing and stereotyping in children. *Child Development, 52*, 1119-1134.

- Meier, S.C., & Labuski, C.M.. (2013). The demographics of the transgender population. In A.K. Baumle (Ed.), *International handbook on the demography of sexuality*. (pp. 289-331). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-5512-3.
- Money, J. (1974). Psychologic consideration of sex assignment in intersexuality. *Clinics in Plastic Surgery*, 1(2), 215–222.
- Money, J., Hampson, J.G., Hampson, J. (1955). An examination of some basic sexual concepts: The evidence of human hermaphroditism. *Bulletin of Johns Hopkins Hospital*, 97(4), 301–319.
- Money, J.,& Ehrhardt, A.A. (1972). *Man and woman, boy and girl: The differentiation and dimorphism of gender identity from conception to maturity*. Baltimore: Johns Hopkins University Press.
- Morris, J. (1974). *Conundrum*. New York, NY: Harcourt Brace Jovanovich.
- Morrison S., & Dinkel S. (2012). Heterosexism and health care: a concept analysis. *Nursing Forum*, 47(2), 123-30.
- Muehlenhard, C. L., & Peterson, Z. D. (2011). Distinguishing Between Sex and Gender: History, Current Conceptualizations, and Implications. *Sex Roles*, 64(11-12), 791–803.  
doi:10.1007/s11199-011-9932-5
- Murphy, M. C., Steele, C. M., & Gross, J. J. (2007). Signaling threat: How situational cues affect women in math, science, and engineering settings. *Psychological Science*, 18(10), 879–885.
- Nagoshi, J., Adams, K., Terrell, H., Hill, E., Brzuzy, S., & Nagoshi, C. (2008). Gender differences in correlates of homophobia and transphobia. *Sex Roles*, 59(7/8), 521-531.  
doi:10.1007/s11199-008-9458-7

- Olson, K. R., Key, A. C., & Eaton, N. R. (2015). Gender Cognition in Transgender Children. *Psychological Science*. doi:10.1177/0956797614568156
- Operario, D., Soma, T., Underhill, K. (2008) Sex work and HIV status among transgender women: Systematic review and meta-analysis. *Journal of Acquired Immune Deficiency Syndromes*. 48(1), 97-103.
- Pope, A.R.D., & Warner, R.H. (2016). Identifying the risk of identity threat in transgender populations. Manuscript submitted for publication.
- Postmes, T., Haslam, S. A., & Jans, L. (2013). A single-item measure of social identification: Reliability, validity, and utility. *British Journal of Social Psychology*, 52(4), 597–617. doi:10.1111/bjso.12006
- Redford, J., & Van Wagenen, A. (2012). *Measuring sexual orientation identity and gender identity in a self-administered survey: Results from cognitive research with older adults*. Retrieved from <http://paa2012.princeton.edu/abstracts/122975>
- Reeve, C.D.C. (2006). *Plato on Love*. Indianapolis, IN: Hackett Publishing Company.
- Reisner, S. L., Conron, K. J., Tardiff, L. A., Jarvi, S., Gordon, A. R., & Austin, S. B. (2014). Monitoring the health of transgender and other gender minority populations: Validity of natal sex and gender identity survey items in a U.S. national cohort of young adults. *BMC Public Health*, 14(1), 1-19. doi:10.1186/1471-2458-14-1224
- Reisner, S.L., Lloyd, J., Baral, S.D. (2013). *Technical Report: The Global Health Needs of Transgender Populations*. USAID’s AIDS Support and Technical Assistance Resources, AIDSTAR-Two, Task Order 2, Arlington, VA.
- Rockquemore, K. A. (1998). Between Black and White: Exploring the “Biracial” Experience. *Race and Society*, 1(2), 197–212.

- Rockquemore, K. A., & Arend, P. (2002). Opting for White: Choice, fluidity and racial identity construction in post civil-rights America. *Race and Society*, 5(1), 49–64.  
<https://doi.org/10.1016/j.racsoc.2003.12.004>
- Roets, A. & Van Hiel, A. (2011). Item selection and validation of a brief, 15-item version of the need for closure scale. *Personality and Individual Differences*, 50, 90-94.
- Rosario, M. & Schrimshaw, E.W. (2014). Theories and etiologies of sexual orientation. In D.L. Tolman & L.M. Diamond (Eds.). *APA Handbook of Sexuality and Psychology: Vol. 1 Person-Based Approaches*. (p. 555- 596). Washington, DC: American Psychological Association. <http://dx.doi.org/10.1037/14193-018>.
- Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton, NJ: Princeton.
- Sanders, S. A., Hill, B. J., Yarber, W. L., Graham, C.A., Crosby, R.A., & Milhausen, R. R. (2010). Misclassification bias: Diversity in conceptualisations about having had sex. *Sexual Health*, 7(1), 31–34. doi:10.1071/SH09068
- Sausa, L. A., Sevelius, J., Keatley, J., Iñiguez, J. R., & Reyes, M. (2009). *Policy recommendations for inclusive data collection of trans people in HIV prevention, care, and services*. Center of Excellence for Transgender HIV Prevention: University of California, San Francisco. Retrieved from [www.transhealth.ucsf.edu](http://www.transhealth.ucsf.edu).
- Schwartz, N. (1999). Cognitive research into survey measurement: Its influence on survey methodology and cognitive theory. In Sirken, M.G., Herrmann, D.J., Schechter, S., Schwartz, N., Tanur, J.M., & Tourangeau, R. (Eds.), *Cognition and Survey Research* (pp. 65-75). New York, NY: John Wiley & Sons, Inc.
- Schwarz, N. (1999). How the Questions Shape the Answers. *American Psychologist*, 54(2), 93-105.

- Scutti, S. (2017). *'The protocol of the day was to lie': NYC issues first US 'intersex' birth certificate*. CNN. Retrieved from <https://www.cnn.com/2016/12/30/health/intersex-birth-certificate/index.html>
- Sell, R. L. (1997). Defining and measuring sexual orientation: a review. *Archives of Sexual Behavior, 26*(6), 643–658.
- Serano, J. M. (2010). The Case Against Autogynephilia. *International Journal of Transgenderism, 12*(3), 176–187. doi:10.1080/15532739.2010.514223
- Serano, J.M. (2007). *Whipping girl: A transsexual woman on sexism and the scapegoating of femininity*. Berkeley, CA: Seal Press.
- Shively, M. G., & DeCecco, J. P. (1977). Components of Sexual Identity. *Journal of Homosexuality, 3*, 41-48.
- Singh, D., Deogracias, J. J., Johnson, L. L., Bradley, S. J., Kibblewhite, S. J., Owen-Anderson, A., ... Zucker, K. J. (2010). The gender identity/gender dysphoria questionnaire for adolescents and adults: further validity evidence. *Journal of Sex Research, 47*(1), 49–58. doi:10.1080/00224490902898728
- Smiler, A.P. & Epstein, M. (2010). Measuring gender: options and issues. In: J.C. Chrisler and D.R. McCreary (Eds.). *Handbook of gender research in psychology: Vol. 1. Gender research in general and experimental psychology*. New York: Springer, 133–157.
- Spence, J. T., & Buckner, C. (1995). Masculinity and femininity: Defining the undefinable. In P. J. Kalbfleisch & M. J. Cody (Eds.), *Gender, power, and communication in human relationships* (pp. 105–138). Hillsdale, NJ: Lawrence Erlbaum Associates



- Staples, J.M., Bird, E.R, Masters, T.N., & George, W.H. (2018). Considerations for culturally sensitive research with transgender adults: A qualitative analysis. *Journal of Sex Research, 55*(8), 1065-1076. <https://doi.org/10.1080/00224499.2017.1292419>
- Stotzer, R. L. (2009). Violence against transgender people: A review of United States data. *Aggression and Violent Behavior, 14*, 170–179.
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. Austin & S. Wochel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
- Tate, C. C., Youssef, C. P., & Bettergarcia, J. N. (2014). Integrating the study of transgender spectrum and cisgender experiences of self-categorization from a personality perspective. *Review of General Psychology, 18*(4), 302–312.
- Tate, C.C., Ledbetter, J.N., & Youssef, C.P. (2013). A two-question method for assessing gender categories in the social and medical sciences. *Journal of Sex Research, 50*(8), 767-776.
- Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology, 78*(5), 853–870.  
<https://doi.org/10.1037//0022-3514.78.5.853>
- Townsend, S. S. M., Markus, H. R., & Bergsieker, H. B. (2009). My choice, your categories: The denial of multiracial identities. *Journal of Social Issues, 65*(1), 185–204.  
[doi:10.1111/j.1540-4560.2008.01594.x](https://doi.org/10.1111/j.1540-4560.2008.01594.x)
- Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15–40). Cambridge: Cambridge University Press.

- van Anders, S. M. (2015). Beyond Sexual Orientation: Integrating Gender/Sex and Diverse Sexualities via Sexual Configurations Theory. *Archives of Sexual Behavior, 44*.  
<http://doi.org/10.1007/s10508-015-0490-8>
- Vrangalova, Z., & Savin-Williams, R. C. (2012). Mostly heterosexual and mostly gay/lesbian: evidence for new sexual orientation identities. *Archives of Sexual Behavior, 41*(1), 85–101.  
[doi:10.1007/s10508-012-9921-y](https://doi.org/10.1007/s10508-012-9921-y)
- Weinberg, M.S., Williams, C.J., & Pryor, D.W. (1994). *Dual attraction: Understanding bisexuality*. Oxford University Press, New York.
- Wright, L. W., Adams, H. E., & Bemat, J. (1999). Development and Validation of the Homophobia Scale, *Journal of Psychopathology & Behavioral Assessment, 21*(4), 1–3.
- Wylie, S. A., Corliss, H. L., Boulanger, V., Prokop, L. A., & Austin, S. B. (2010). Socially assigned gender nonconformity: A brief measure for use in surveillance and investigation of health disparities. *Sex Roles, 63*(3), 264–276. [doi:10.1007/s11199-010-9798-y](https://doi.org/10.1007/s11199-010-9798-y)
- Xavier, J., Honnold, J.A., Bradford, J. (2007). *The Health, Health-Related Needs, and Lifecourse Experiences of Transgender Virginians*. Richmond, VA: Virginia Department of Health, Division of Disease Prevention.
- Zimmerman, J. (2014). Not All Men: A Brief History of Every Dude's Favorite Argument. *TIME*. Retrieved from <http://time.com/79357/not-all-men-a-brief-history-of-every-dudes-favorite-argument/>
- Zucker, K. J. (2005). Gender identity disorder in children and adolescents. *Annual Review of Clinical Psychology, 1*, 467–492.

Zucker, K. J., Mitchell, J. N., Bradley, S. J., Tkachuk, J., Cantor, J. M., & Allin, S. M. (2006).

The recalled childhood gender identity/gender role questionnaire: Psychometric properties.

*Sex Roles*, 54(7-8), 469–483. doi:10.1007/s11199-006-9019-x

Zucker, K. J., Wood, H., Singh, D., & Bradley, S. J. (2012). A developmental, biopsychosocial

model for the treatment of children with gender identity disorder. *Journal of*

*Homosexuality*, 59, 369–397.

## Appendix A

### Summary of Materials for Study 1 and Study 2

#### Study 1 – Spring 2016 (no randomization).

*Scales in italics were only included in the gender minority sample*

- Informed consent
- One randomly assigned context manipulation
  - Healthcare/Psych/Dating Website
- One randomly assigned gender question
  - Binary + Other, Multiple-Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity
- Gender salience
- Free text comments about gender question
- Identity threat dependent variables
- Perceptions of researchers
- Attitudes toward researchers
- Emotions toward researchers
- Gender salience, safety, and motivation to be perceived accurately by context (Dating Website, Clinic Intake, Psych Experiment)
- Perceptions of each gender question
  - Binary, Binary + Other, Multiple- Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Multiple-Choice + Transition History, Short Free-Text Box, Long Free-Text Box
- Ranking of gender questions
  - Binary, Binary + Other, Multiple- Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Multiple-Choice + Transition History, Short Free-Text Box, Long Free-Text Box
- *Gender Centrality (adapted from Leach et al., 2008)*
- *Single-Item Social Identification Scales (adapted from Postmes et al., 2013)*
- Gender self-categorization (Joel et al., 2014)
- Multi-Component Social Identity Scale (adapted from Leach et al., 2008)
- Emotions toward gender groups
  - Others with same identity, trans men, trans women, cisgender women, cisgender men, gender non-conforming people, people with non-binary gender identities. Free-text box for comments about these questions.
- Demographics
  - Age
  - Race/ethnicity
  - Area from
  - Personal SES
  - Highest level of education
  - Anything else you would like us to know about you?
- Thank you and raffle link

## **Study 2 – Fall 2016.**

- Informed consent
- One randomly assigned gender question
  - Binary, Binary + Other, Multiple- Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Short Free-Text Box
- Free text comments
- Immediate DVs (randomized order in section and within scales)
  - Evaluation of question
  - Authenticity (Kraus, Chen, and Keltner, 2011)
  - Collective Self-Esteem (Luhtanen & Crocker, 1992)
- Secondary DVs (randomized order in section and within scales, each scale preceded by a reminder of the question randomly assigned at the beginning of study)
  - Expectations of belonging and respect
  - Perceptions of researcher knowledge/concern about transgender people
  - Comfort with/investment in research
  - Emotions toward researchers
- Perceptions of gender questions (randomized order within section)
  - Binary, Binary + Other, Multiple- Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Short Free-Text Box
- Rankings of gender questions by context (randomized order of contexts in section and of gender questions within each context)
  - Clinic Intake, Dating Website, Psych Experiment, Job Application
  - Binary, Binary + Other, Multiple- Choice, Multiple-Choice + Sex Assigned at Birth, Multiple-Choice + Transgender Identity, Short Free-Text Box, No gender question asked
- Supplemental Questions (randomized order in section and within scales)
  - Gender self-categorization (Joel et al., 2014)
  - Gender self-categorization (adapted from Rockquemore & Arend, 2002)
  - Non-binary experiences
  - Emotions toward own gender group
  - Gender Identification (Egan & Perry, 2001)
  - Transgender Congruence Scale (Kozee et al., 2012)
  - Gender Identity Reflection and Rumination scale (Bauerband & Galupo, 2014)
  - Gender essentialism
  - Multi-Gender Identity Questionnaire (Joel et al., 2014)
- Demographic Questions (randomized within section)
  - Political Affiliation, Choice for president in 2016
  - Age, Race/ethnicity, Highest level of education
- Thank you and raffle link

## Coding Scheme for Free-Text Gender Questions in Spring 2016 and Fall 2016 Surveys

### *Gender Identity*

1. Participant identifies with a non-binary identity: the participant lists words like androgynous, bigender, demiboy/demigirl, non-binary, genderfluid, genderqueer, genderfucker, neutrois, queer, or a mixture of these with other gender categories.  
**Preference is given for non-binary identities in descriptions such as “genderqueer woman” or “agender non-binary.”**
2. The participant identifies as agender: the participant identifies as agender or has no gender (none). If the participant lists “agender” with other genders, preference is given to “agender”. \* **Preference is given for agender in descriptions such as “agender woman.”**
3. Participant identifies as a man: the participant lists words like man, male, masculine, boy, transman, AMAB/DMAB, cis-male, or cis-man; there are not other secondary gender categories like in “genderfluid man” or “masculine genderqueer”.
4. Participant identifies as a woman: the participant lists words like woman, female, femme, girl, transwoman, AFAB/DFAB, cis-woman, or cis-female; there are not secondary gender categories like in “genderfluid woman” or “nonbinary woman”.
5. Participant does not know their gender: the participant writes in “I don’t know” or “not sure” into a category. Unknown can also go into this category.
6. Participant rejects the construct of gender: the participant says “gender is made up”, “no thanks”, does not identify with gender roles, or does not have a gender identity (this is different from saying ‘agender’). This can also apply to people who say that their gender is irrelevant. \*

7. Participant does not provide enough information to determine a gender: There is not enough information provided to determine a gender category, descriptions such as human, weariness, cyborg (or other inanimate objects, unless they're an internet meme), me, gendermeh, gender non-conforming (without any other gender category), anything that doesn't seem like it fits into other categories.

\* A person who is agender has no gender – a person who has no gender identity is rejecting the concept of identifying with any gender that they know of.

### ***Trans Status***

1. The participant is definitely cis (cis-woman, AFAB/DFAB woman).
2. The participant is binary trans (trans-woman, trans-man, male-bodied woman).
3. The participant is non-binary (any non-binary identity, non-binary man).
4. The participant's trans status is uncertain (researcher cannot tell if the person is trans with provided information – woman, man).
5. The participant refused to say if they are trans (I refuse to say, it doesn't matter).

### ***Gender Nonconforming***

- Yes – the participant says they are gender non-conforming or do not perform gender in a stereotypical manner.
- No – the participant does not mention being gender non-conforming.

### ***Body mentioned?***

- Yes – the participant mentions their body or genitals without being asked; say their assigned sex at birth (AMAB, AFAB, DMAB, DFAB), say \_\_\_\_\_-bodied, or mention physical transition (transsexual).
- No – the participant does not mention their body.

***Gender expression mentioned?***

- Yes – the participant mentions a specific gender expression (androgynous, femme, butch, masculine-presenting, or queer-presenting), says “I look like \_\_\_\_\_”, or “I am seen as \_\_\_\_\_”.
- No – the participant does not mention a specific gender expression.

***Age-based gender identity?***

- Yes – the participant says they are a boy/girl rather than a man/woman, or mentions in comments that they are definitely a man/woman but not a boy/girl.
- No – the participant does not make any comments about the age.

***Multiple genders?***

- Yes – the participant lists multiple genders in the category (genderqueer woman, agender non-binary, queer transman).
- No – the participant only lists one gender, possibly with a modifier (trans-woman, gender non-conforming man).

***Changing or fluid gender?***

- Yes – the participant mentions being genderfluid or that their gender changes.
- No – the participant does not mention fluid or changing gender.

***Less conventional gender?***



- Yes – the participant lists an unconventional non-binary gender (i.e. cyborg, me, gendermeh, etc.).
- No – the participant uses a listed gender descriptor in the man, woman, agender, or non-binary category.

***Transphobic/Homophobic***

- Yes – The participant writes in transphobic/homophobic commentary, slurs, or terms commonly associated with transphobic or homophobic memes (e.g. doodookin, attack helicopter).

## Appendix B

Table B1

*T-tests for Significant Interactions of Context and Transgender Identity on Dependent Variables (Study 1)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Perceptions of Researchers' Knowledge/Caring</b>					
(Intercept)	4.80		0.14	32.66	<.001
Dating Context		-0.59	0.19	-3.05	.002
Experimental Context		-0.14	0.19	-0.73	<i>ns</i>
Transgender		0.02	0.20	0.13	<i>ns</i>
Dating Context * Transgender		0.77	0.29	2.64	.008
Experiment Context * Transgender		0.55	0.29	1.90	<i>ns</i>

*ns p* >.05

Table B2

*T-tests for Significant Main Effects of Context on Dependent Variables (Study 1).*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Identity Safety</b>					
Clinic Intake	5.32		0.09	55.53	<.001
Dating Website		-0.68	0.13	-5.16	<.001
Psychology Experiment		0.24	0.13	1.84	<i>ns</i>
<b>Perceptions of Researchers' Knowledge/Caring</b>					
Clinic Intake	4.74		0.10	47.40	<.001
Dating Website		-0.29	0.13	-2.11	.034
Psychology Experiment		0.11	0.13	0.85	<i>ns</i>

*ns p* >.05

Table B3

*T-tests for Significant Main Effects of Transgender Identity on Dependent Variables (Study 1).*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Identity Safety</b>					
Not transgender	5.69		0.08	64.39	<.001
Transgender		-0.80	0.11	-6.86	<.001
Unclassified		-0.63	0.33	-1.90	<i>ns</i>
<b>Perceptions of Researchers' Knowledge/Caring</b>					
Not transgender	5.02		0.08	56.08	<.001
Transgender		-0.48	0.11	-4.073	<.001
Unclassified		-0.50	0.33	-1.50	<i>ns</i>
<b>Attitudes Toward Researchers</b>					
Not transgender	2.70		0.08	30.75	<.001
Transgender		0.35	0.11	3.00	.002
Unclassified		0.39	0.32	1.22	<i>ns</i>
<b>Positive Emotions Toward Researchers</b>					
Not transgender	3.25		0.11	28.59	<.001
Transgender		0.39	0.15	2.62	.008
Unclassified		0.09	0.42	0.21	<i>ns</i>

*ns p* >.05

Table B4

*T*-tests for Significant Interactions of Question Condition and Transgender Identity on Dependent Variables (Study 2)

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Comfort with Research</b>					
Intercept	4.72		0.10	43.80	<.001
Binary + Other		0.07	0.15	0.49	<i>ns</i>
Multiple-Choice		0.35	0.15	2.24	.02
MC + Sex Assigned at Birth		0.43	0.16	2.67	.007
MC + Transgender Identity		0.41	0.16	2.58	.01
Short Free Text		0.49	0.17	2.90	.003
Transgender		-1.82	0.39	-4.66	<.001
Unclassified		-0.50	0.71	-0.71	<i>ns</i>
Binary + Other * Transgender		1.99	0.48	4.15	<.001
Multiple-Choice * Transgender		2.27	0.57	3.94	<.001
MC + Sex Assigned at Birth * Transgender		1.77	0.55	3.19	.001
MC + Transgender Identity * Transgender		1.97	0.51	3.82	<.001
Short Free Text * Transgender		2.04	0.49	4.11	<.001
Binary + Other * Unclassified		-0.37	0.84	-0.44	<i>ns</i>
Multiple-Choice * Unclassified		0.24	0.81	0.30	<i>ns</i>
MC + Sex Assigned at Birth * Unclassified		-0.71	1.00	-0.71	<i>ns</i>
MC + Transgender Identity * Unclassified		-1.29	0.92	-1.41	<i>ns</i>
Short Free Text * Unclassified		-1.04	0.92	-1.13	<i>ns</i>

*ns p* >.05

Note: Reference categories are the binary question and cisgender participants.

Table B5

*T-tests for Significant Main Effects of Question Conditions on Dependent Variables (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Expectation of Belonging/Respect</b>					
Binary	5.04		0.13	37.49	<.001
Binary + Other		0.27	0.18	1.43	<i>ns</i>
Multiple-Choice		0.85	0.19	4.41	<.001
MC + Sex Assigned at Birth		0.68	0.19	3.46	<.001
MC + Transgender Identity		0.69	0.19	3.62	<.001
Short Free Text		0.54	0.20	2.67	.007
<b>Comfort with Research</b>					
Binary	4.54		0.10	42.16	<.001
Binary + Other		0.24	0.15	1.62	<i>ns</i>
Multiple-Choice		0.47	0.15	3.07	.002
MC + Sex Assigned at Birth		0.54	0.15	3.38	<.001
MC + Transgender Identity		0.55	0.15	3.63	<.001
Short Free Text		0.66	0.16	4.09	<.001
<b>Perceptions of Researchers</b>					
Binary	3.46		0.11	30.58	<.001
Binary + Other		0.79	0.15	5.04	<.001
Multiple-Choice		1.57	0.16	9.69	<.001
MC + Sex Assigned at Birth		1.65	0.16	9.88	<.001
MC + Transgender Identity		1.50	0.16	9.36	<.001
Short Free Text		1.40	0.17	8.21	<.001
<b>Positive Emotions Toward Researchers</b>					
Binary	2.55		0.16	15.14	<.001
Binary + Other		0.41	0.23	1.73	<i>ns</i>
Multiple-Choice		1.23	0.24	5.09	<.001
MC + Sex Assigned at Birth		1.36	0.24	5.48	<.001
MC + Transgender Identity		0.74	0.23	3.10	.002
Short Free Text		1.12	0.25	4.39	<.001
<b>Negative Emotions Toward Researchers</b>					
Binary	2.58		0.13	19.306	<.001
Binary + Other		-0.54	0.18	-2.94	.003
Multiple-Choice		-0.57	0.19	-2.98	.002
MC + Sex Assigned at Birth		-0.72	0.19	-3.70	<.001
MC + Transgender Identity		-0.74	0.18	-3.95	<.001
Short Free Text		-0.45	0.20	-2.22	.026

*ns p >.05*

Table B6

*T-tests for Significant Main Effects of Transgender Identity on Dependent Variables (Study 2).*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Evaluation of Question</b>					
Not transgender	4.44		0.03	134.46	<.001
Transgender		-0.59	0.10	-5.96	<.001
Unclassified		-0.62	0.15	-4.13	<.001
<b>Authenticity</b>					
Not transgender	5.32		0.05	104.64	<.001
Transgender		-0.97	0.15	-6.31	<.001
Unclassified		-1.27	0.23	-5.44	<.001
<b>CSES (Private Subscale)</b>					
Not transgender	5.75		0.04	142.05	<.001
Transgender		-0.29	0.12	-2.40	.016
Unclassified		-1.30	0.18	-6.98	<.001
<b>CSES (Public Subscale)</b>					
Not transgender	5.04		0.05	91.42	<.001
Transgender		-1.54	0.16	-9.33	<.001
Unclassified		-1.24	0.25	-4.83	<.001
<b>Expectation of Belonging/Respect</b>					
Not transgender	5.77		0.05	100.05	<.001
Transgender		-1.05	0.17	-6.08	<.001
Unclassified		-1.24	0.26	-4.70	<.001
<b>Comfort with Research</b>					
Not transgender	5.02		0.04	107.87	<.001
Transgender		-0.17	0.14	-1.28	<i>ns</i>
Unclassified		-0.83	0.22	-3.73	<.001
<b>Perceptions of Researchers</b>					
Not transgender	4.59		0.05	82.43	<.001
Transgender		0.19	0.16	1.16	<i>ns</i>
Unclassified		-0.30	0.25	-1.18	<i>ns</i>
<b>Positive Emotions Toward Researchers</b>					
Not transgender	3.23		0.07	42.32	<.001
Transgender		0.66	0.23	2.84	.004
Unclassified		0.50	0.34	1.44	<i>ns</i>
<b>Negative Emotions Toward Researchers</b>					
Not transgender	2.02		0.05	34.24	<.001
Transgender		0.14	0.17	0.81	<i>ns</i>
Unclassified		0.62	0.27	2.26	.024

*ns p* >.05

Table B7

*T*-tests for Significant Main Effects of Gender Identity on Evaluations of Gender Questions  
(Study 1)

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	5.68		0.13	43.01	<.001
Woman		-0.40	0.18	-2.12	.03
Nonbinary/Agender		-3.90	0.18	-20.67	<.001
<b>Binary + Other</b>					
Man	5.92		0.12	47.08	<.001
Woman		-0.18	0.17	-1.06	<i>ns</i>
Nonbinary/Agender		-0.96	0.17	-5.40	<.001
<b>Multiple-Choice</b>					
Man	5.98		0.11	51.65	<.001
Woman		-0.10	0.16	-0.66	<i>ns</i>
Nonbinary/Agender		-0.47	0.16	-2.85	.004
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Man	5.96		0.11	50.57	<.001
Woman		0.15	0.16	0.91	<i>ns</i>
Nonbinary/Agender		-0.85	0.16	-5.069	<.001
<b>Multiple-Choice + Transgender Identity</b>					
Man	5.90		0.11	50.91	<.001
Woman		0.13	0.16	0.81	<i>ns</i>
Nonbinary/Agender		-0.41	0.16	-2.48	.01
<b>Multiple-Choice + Transition History</b>					
Man	5.84		0.13	44.26	<.001
Woman		-0.08	0.18	-0.47	<i>ns</i>
Nonbinary/Agender		-1.29	0.18	-6.87	<.001
<b>Short Free-Text</b>					
Man	6.33		0.09	65.67	<.001
Woman		0.03	0.13	0.26	<i>ns</i>
Nonbinary/Agender		-0.46	0.13	-3.40	<.001
<b>Long Free-Text</b>					
Man	6.37		0.09	69.73	<.001
Woman		0.10	0.13	0.77	<i>ns</i>
Nonbinary/Agender		0.02	0.13	0.17	<i>ns</i>

*ns p* >.05



Table B8

*T-tests for Significant Main Effects of Transgender Identity on Perceptions of Gender Questions (Study 1)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	5.98		0.12	48.66	<.001
Transgender		-3.33	0.16	-20.37	<.001
Unclassified		-2.33	0.35	-6.62	<.001
<b>Binary + Other</b>					
Not transgender	6.17		0.10	59.27	<.001
Transgender		-1.19	0.13	-8.64	<.001
Unclassified		-2.12	0.30	-7.00	<.001
<b>Multiple-Choice</b>					
Not transgender	6.23		0.09	64.10	<.001
Transgender		-0.85	0.12	-6.57	<.001
Unclassified		-1.73	0.28	-6.10	<.001
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Not transgender	6.48		0.09	67.66	<.001
Transgender		-1.43	0.12	-11.24	<.001
Unclassified		-1.88	0.27	-6.73	<.001
<b>Multiple-Choice + Transgender Identity</b>					
Not transgender	6.30		0.09	66.12	<.001
Transgender		-0.82	0.12	-6.49	<.001
Unclassified		-2.33	0.27	-8.40	<.001
<b>Multiple-Choice + Transition History</b>					
Not transgender	6.26		0.10	58.02	<.001
Transgender		-1.61	0.14	-11.23	<.001
Unclassified		-2.28	0.31	-7.25	<.001
<b>Short Free-Text</b>					
Not transgender	6.47		0.08	79.18	<.001
Transgender		-0.51	0.10	-4.71	<.001
Unclassified		-1.15	0.23	-4.84	<.001
<b>Long Free-Text</b>					
Not transgender	6.53		0.07	83.26	<.001
Transgender		-0.18	0.10	-1.73	<i>ns</i>
Unclassified		-1.15	0.23	-4.95	<.001

*ns p* >.05

Table B9

*T-tests for Significant Interactions of Gender Identity and Transgender Identity on Rankings of Gender Questions (Study 1)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Multiple-Choice + Assigned Sex at Birth</b>					
(Intercept)	4.19		0.19	21.35	<.001
Woman		-0.02	0.28	-0.10	<i>ns</i>
Non-binary/Agender		-3.19	1.78	-1.79	<i>ns</i>
Transgender		1.02	0.35	2.89	.004
Unclassified		-1.03	0.74	-1.37	<i>ns</i>
Woman * Transgender		-0.69	0.49	-1.40	<i>ns</i>
Non-binary/Agender * Transgender		2.91	1.81	1.60	<i>ns</i>
Woman * Unclassified		3.06	1.10	2.76	.006
Non-binary/Agender * Unclassified		4.78	2.11	2.26	.024
<b>Short Free-Text</b>					
(Intercept)	4.24		0.24	17.66	<.001
Woman		-0.13	0.34	-0.37	<i>ns</i>
Non-binary/Agender		1.75	2.17	0.80	<i>ns</i>
Transgender		-1.99	0.43	-4.60	<.001
Unclassified		1.41	0.91	1.55	<i>ns</i>
Woman * Transgender		1.25	0.60	2.07	.038
Non-binary/Agender * Transgender		-1.15	2.21	-0.52	<i>ns</i>
Woman * Unclassified		-3.73	1.35	-2.75	.006
Non-binary/Agender * Unclassified		-4.66	2.58	-1.80	<i>ns</i>
<b>Long Free-Text</b>					
(Intercept)	4.49		0.29	15.28	<.001
Woman		-0.29	0.42	-0.71	<i>ns</i>
Non-binary/Agender		3.50	2.66	1.31	<i>ns</i>
Transgender		-0.88	0.53	-1.66	<i>ns</i>
Unclassified		2.67	1.11	2.38	.017
Woman * Transgender		-0.08	0.73	-0.11	<i>ns</i>
Non-binary/Agender * Transgender		-4.50	2.70	-1.66	<i>ns</i>
Woman * Unclassified		-5.66	1.65	-3.42	<.001
Non-binary/Agender * Unclassified		-9.17	3.16	-2.89	.003

*Note.* Higher numbers indicate lower ranking (i.e. 8 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

Table B10

*T*-tests for Significant Main Effects of Gender Identity on Rankings of Gender Questions (Study 1).

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	5.14		0.21	23.73	<.001
Woman		0.81	0.30	2.64	.008
Nonbinary/Agender		2.54	0.30	8.41	<.001
<b>Binary + Other</b>					
Man	4.14		0.18	22.59	<.001
Woman		0.41	0.26	1.58	<i>ns</i>
Nonbinary/Agender		1.18	0.25	4.65	<.001
<b>Multiple-Choice</b>					
Man	3.99		0.13	28.74	<.001
Woman		-0.14	0.19	-0.75	<i>ns</i>
Nonbinary/Agender		-0.57	0.19	-2.99	.001
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Man	4.44		0.16	27.27	<.001
Woman		-0.08	0.23	-0.37	<i>ns</i>
Nonbinary/Agender		0.46	0.22	2.03	.04
<b>Multiple-Choice + Transgender Identity</b>					
Man	4.60		0.13	32.95	<.001
Woman		-0.26	0.19	-1.34	<i>ns</i>
Nonbinary/Agender		-0.86	0.19	-4.46	<.001
<b>Short Free-Text</b>					
Man	3.73		0.20	18.42	<.001
Woman		0.04	0.28	0.16	<i>ns</i>
Nonbinary/Agender		-0.86	0.28	-3.08	.002
<b>Long Free-Text</b>					
Man	4.36		0.24	17.90	<.001
Woman		-0.61	0.34	-1.77	<i>ns</i>
Nonbinary/Agender		-1.73	0.33	-5.12	<.001

*Note.* Higher numbers indicate lower ranking (i.e. 8 is the lowest possible ranking, and 1 is the highest).

*ns*  $p > .05$

Table B11

*T-tests for Significant Main Effects of Transgender Identity on Rankings of Gender Questions (Study 1)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	5.07		0.18	28.00	<.001
Transgender		2.28	0.23	9.76	<.001
Unclassified		1.09	0.50	2.17	.030
<b>Binary + Other</b>					
Not transgender	4.11		0.15	26.05	<.001
Transgender		1.12	0.20	5.49	<.001
Unclassified		0.17	0.43	0.41	<i>ns</i>
<b>Multiple-Choice</b>					
Not transgender	3.99		0.12	32.52	<.001
Transgender		-0.39	0.15	-2.51	.012
Unclassified		-0.32	0.34	-0.96	<i>ns</i>
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Not transgender	4.15		0.14	29.30	<.001
Transgender		0.82	0.18	4.49	<.001
Unclassified		-0.11	0.39	-0.28	<i>ns</i>
<b>Multiple-Choice + Transgender Identity</b>					
Not transgender	4.65		0.12	38.75	<.001
Transgender		-0.87	0.15	-5.60	<.001
Unclassified		0.04	0.33	0.15	<i>ns</i>
<b>Short Free-Text</b>					
Not transgender	4.16		0.17	24.37	<.001
Transgender		-1.31	0.22	-5.96	<.001
Unclassified		-0.41	0.47	-0.86	<i>ns</i>
<b>Long Free-Text</b>					
Not transgender	4.39		0.20	20.98	<.001
Transgender		-1.61	0.27	-5.94	<.001
Unclassified		-0.43	0.58	-0.75	<i>ns</i>

*Note.* Higher numbers indicate lower ranking (i.e. 8 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

Table B12

*T-tests for Significant Main Effects of Gender Identity on Perceptions of Gender Questions (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	4.41		0.05	79.52	<.001
Woman		-0.20	0.07	-2.70	.007
Nonbinary/Agender		-1.16	0.12	-9.00	<.001
<b>Binary + Other</b>					
Man	4.50		0.05	79.81	<.001
Woman		-0.01	0.07	-0.13	<i>ns</i>
Nonbinary/Agender		-1.25	0.13	-9.36	<.001
<b>Multiple-Choice</b>					
Man	4.53		0.05	89.37	<.001
Woman		0.24	0.06	3.48	<.001
Nonbinary/Agender		-0.53	0.11	-4.54	<.001
<b>MC + Assigned Sex at Birth</b>					
Man	4.42		0.05	79.02	<.001
Woman		0.23	0.07	3.07	.002
Nonbinary/Agender		-0.89	0.13	-6.88	<.001
<b>MC + Transgender Identity</b>					
Man	4.39		0.05	86.39	<.001
Woman		0.22	0.06	3.28	.001
Nonbinary/Agender		-0.35	0.11	-3.03	.002
<b>Short Free-Text</b>					
Man	4.67		0.05	90.74	<.001
Woman		0.16	0.07	2.40	.016
Nonbinary/Agender		-0.58	0.12	-4.88	<.001

*ns p* >.05

Table B13

*T-tests for Significant Main Effects of Transgender Identity on Evaluations of Gender Questions (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	4.32		0.03	109.70	<.001
Transgender		-0.94	0.11	-7.97	<.001
Unclassified		-0.41	0.18	-2.29	.021
<b>Binary + Other</b>					
Not transgender	4.50		0.03	113.37	<.001
Transgender		-1.02	0.12	-8.50	<.001
Unclassified		-0.51	0.18	-2.85	.004
<b>Multiple-Choice</b>					
Not transgender	4.67		0.03	131.14	<.001
Transgender		-0.56	0.10	-5.29	<.001
Unclassified		-0.39	0.16	-2.43	.015
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Not transgender	4.57		0.03	118.21	<.001
Transgender		-1.03	0.11	-8.91	<.001
Unclassified		-0.26	0.17	-1.53	<i>ns</i>
<b>Multiple-Choice + Transgender Identity</b>					
Not transgender	4.53		0.03	127.12	<.001
Transgender		-0.49	0.10	-4.56	<.001
Unclassified		-0.59	0.16	-3.60	<.001
<b>Short Free-Text</b>					
Not transgender	4.78		0.03	134.83	<.001
Transgender		-0.72	0.10	-6.76	<.001
Unclassified		-0.41	0.16	-2.48	.013

*ns p* >.05

Table B14

*T-tests for Significant Main Effects of Gender Identity on Rankings of Gender Questions in the Clinic Intake Context (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	4.90		0.13	36.66	<.001
Woman		0.86	0.18	4.74	<.001
Nonbinary/Agender		1.86	0.29	6.33	<.001
<b>Binary + Other</b>					
Man	3.70		0.11	31.78	<.001
Woman		0.67	0.15	4.25	<.001
Nonbinary/Agender		1.35	0.25	5.25	<.001
<b>Multiple-Choice</b>					
Man	3.59		0.09	36.97	<.001
Woman		-0.36	0.13	-2.77	.005
Nonbinary/Agender		-0.67	0.21	-3.14	.001
<b>Multiple-Choice + Transgender Identity</b>					
Man	4.02		0.11	34.62	<.001
Woman		-0.43	0.15	-2.70	.006
Nonbinary/Agender		-1.18	0.25	-4.61	<.001

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

Table B15

*T*-tests for Significant Main Effects of Gender Identity on Rankings of Gender Questions in the Psychology Experiment Context (Study 2)

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	4.91		0.13	35.15	<.001
Woman		0.81	0.18	4.30	<.001
Nonbinary/Agender		1.88	0.30	6.21	<.001
<b>Binary + Other</b>					
Man	3.72		0.12	30.95	<.001
Woman		0.71	0.16	4.39	<.001
Nonbinary/Agender		1.09	0.26	4.18	<.001
<b>Multiple-Choice</b>					
Man	3.84		0.10	37.36	<.001
Woman		-0.41	0.13	-2.98	.002
Nonbinary/Agender		-0.94	0.22	-4.22	<.001
<b>Multiple-Choice + Transgender Identity</b>					
Man	3.81		0.12	31.04	<.001
Woman		-0.42	0.16	-2.54	.011
Nonbinary/Agender		-1.09	0.26	-4.0	<.001

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05



Table B16

*T*-tests for Significant Main Effects of Gender Identity on Rankings of Gender Questions in the Dating Website Context (Study 2)

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	4.33		0.14	30.84	<.001
Woman		0.87	0.18	4.66	<.001
Nonbinary/Agender		2.28	0.30	7.40	<.001
<b>Binary + Other</b>					
Man	3.29		0.11	28.58	<.001
Woman		0.62	0.15	4.04	<.001
Nonbinary/Agender		0.78	0.25	3.11	.001
<b>Multiple-Choice</b>					
Man	3.86		0.10	35.58	<.001
Woman		-0.47	0.14	-3.25	.001
Nonbinary/Agender		-1.33	0.23	-5.58	<.001
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Man	3.94		0.13	28.80	<.001
Woman		-0.39	0.18	-2.16	.031
Nonbinary/Agender		1.05	0.30	3.52	<.001
<b>Multiple-Choice + Transgender Identity</b>					
Man	3.96		0.12	31.99	<.001
Woman		-0.38	0.16	-2.33	.019
Nonbinary/Agender		-0.53	0.27	-1.96	.050
<b>Short Free-Text</b>					
Man	3.09		0.12	24.14	<.001
Woman		-0.46	0.17	-2.71	.006
Nonbinary/Agender		-1.32	0.28	-4.69	<.001
<b>No Question Asked</b>					
Man	5.51		0.13	39.415	<.001
Woman		0.22	0.18	1.18	<i>ns</i>
Nonbinary/Agender		-0.94	0.30	-3.06	.002

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

Table B17

*T*-tests for Significant Main Effects of Gender Identity on Rankings of Gender Questions in the Job Application Context (Study 2)

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Man	4.69		0.15	30.35	<.001
Woman		0.57	0.20	2.79	.005
Nonbinary/Agender		1.78	0.33	5.37	<.001
<b>Binary + Other</b>					
Man	3.29		0.12	26.96	<.001
Woman		0.58	0.16	3.62	<.001
Nonbinary/Agender		0.64	0.26	2.46	.014
<b>Multiple-Choice</b>					
Man	3.95		0.11	34.76	<.001
Woman		-0.30	0.15	-2.02	.043
Nonbinary/Agender		-0.97	0.24	-4.00	<.001
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Man	4.35		0.12	35.10	<.001
Woman		0.21	0.16	1.30	<i>ns</i>
Nonbinary/Agender		1.16	0.26	4.37	<.001
<b>Short Free-Text</b>					
Man	3.34		0.12	26.97	<.001
Woman		-0.50	0.16	-3.08	.002
Nonbinary/Agender		-0.88	0.26	-3.33	<.001
<b>No Question Asked</b>					
Man	4.00		0.18	22.09	<.001
Woman		-0.85	0.24	-3.55	<.001
Nonbinary/Agender		-1.84	0.38	-4.75	<.001

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns*  $p > .05$

Table B18

*T-tests for Significant Main Effects of Transgender Identity on Rankings of Gender Questions in the Clinic Intake Context (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	5.38		0.09	56.11	<.001
Transgender		1.11	0.27	4.07	<.001
Unclassified		0.11	0.44	0.26	<i>ns</i>
<b>Binary + Other</b>					
Not transgender	4.07		0.08	49.03	<.001
Transgender		0.62	0.23	2.64	.008
Unclassified		0.97	0.38	2.50	.012
<b>Multiple-Choice</b>					
Not transgender	3.42		0.06	50.62	<.001
Transgender		-0.40	0.19	-2.09	.036
Unclassified		-0.82	0.31	-2.60	.009
<b>Multiple-Choice + Transgender Identity</b>					
Not transgender	3.80		0.08	46.57	<.001
Transgender		-0.78	0.23	-3.38	<.001
Unclassified		-0.30	0.38	-0.80	<i>ns</i>

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns*  $p > .05$

Table B19

*T-tests for Significant Main Effects of Transgender Identity on Rankings of Gender Questions in the Psychology Experiment Context (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	5.36		0.09	54.42	<.001
Transgender		1.17	0.27	4.22	<.001
Unclassified		0.33	0.44	0.76	<i>ns</i>
<b>Multiple-Choice + Transgender Identity</b>					
Not transgender	3.57		0.08	41.70	<.001
Transgender		-0.72	0.24	-2.97	.003
Unclassified		-0.07	0.38	-0.20	<i>ns</i>

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

Table B20

*T*-tests for Significant Main Effects of Transgender Identity on Rankings of Gender Questions in the Dating Website Context (Study 2)

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	4.82		0.90	49.19	<.001
Transgender		1.36	0.28	4.79	<.001
Unclassified		0.99	0.50	1.96	.049
<b>Multiple-Choice</b>					
Not transgender	3.59		0.07	48.60	<.001
Transgender		-1.03	0.21	-4.82	<.001
Unclassified		-0.53	0.38	-1.41	<i>ns</i>
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Not transgender	3.71		0.09	39.64	<.001
Transgender		1.08	0.27	3.97	<.001
Unclassified		0.53	0.48	1.09	<i>ns</i>
<b>Short Free-Text</b>					
Not transgender	2.82		0.08	31.75	<.001
Transgender		-0.54	0.25	-2.12	.033
Unclassified		-0.75	0.45	-1.66	<i>ns</i>
<b>No Question Asked</b>					
Not transgender	5.62		0.09	58.87	<.001
Transgender		-0.82	0.27	-2.99	.002
Unclassified		-0.50	0.49	-1.02	<i>ns</i>

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

Table B21

*T-tests for Significant Main Effects of Transgender Identity on Rankings of Gender Questions in the Job Application Context (Study 2)*

	Mean	Mean difference from reference	SE	<i>t</i>	<i>p</i>
<b>Binary</b>					
Not transgender	5.01		0.10	48.08	<.001
Transgender		1.28	0.30	4.22	<.001
Unclassified		0.87	0.48	1.79	<i>ns</i>
<b>Multiple-Choice</b>					
Not transgender	3.79		0.07	49.53	<.001
Transgender		-0.72	0.22	-3.22	.001
Unclassified		-0.58	0.35	-1.63	<i>ns</i>
<b>Multiple-Choice + Assigned Sex at Birth</b>					
Not transgender	4.44		0.08	53.01	<.001
Transgender		1.12	0.24	4.58	<.001
Unclassified		0.18	0.39	0.47	<i>ns</i>
<b>No Question Asked</b>					
Not transgender	3.58		0.12	29.31	<.001
Transgender		-1.62	0.35	-4.54	<.001
Unclassified		-0.63	0.57	-1.11	<i>ns</i>

*Note.* Higher numbers indicate lower ranking (i.e. 7 is the lowest possible ranking, and 1 is the highest).

*ns p* >.05

## Appendix C

Table C1

*Cronbach's Alpha, Means, and Standard Deviations for Emotions Toward Gender Groups (Studies 1 and 2).*

Study	Gender Group	Positive Emotions			Negative Emotions		
		$\alpha$	M	SD	$\alpha$	M	SD
1	People with the same gender identity as you	.87	5.0	1.5	.93	2.4	1.4
	Transgender men	.92	4.5	1.8	.93	2.0	1.2
	Transgender women	.92	4.7	1.9	.93	1.9	1.2
	Cisgender men	.91	4.1	1.7	.93	2.3	1.4
	Cisgender women	.88	3.3	1.6	.93	2.9	1.7
	Gender non-conforming people	.93	4.6	1.9	.96	1.9	1.2
	People with non-binary gender identities	.94	4.6	2.0	.96	1.9	1.2
2	People with the same gender identity as you	.89	3.3	1.7	.92	2.1	1.3

Table C2

*Fixed-Effects ANOVAs for Interactions of Gender Identity and Transgender Identity on Identity as a Man (Multi-Dimensional Social Identity Scale, Leach et al. 2008) - Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2458.62	1	2458.62	1515.82	.000		
summgen3	30.17	2	15.09	9.30	.000	.05	[.02, .08]
transid3	17.33	2	8.66	5.34	.005	.03	[.00, .06]
summgen3 x transid3	19.50	4	4.88	3.01	.018	.03	[.00, .05]
Error	624.46	385	1.62				

Note. LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C3

*Expected Means for Gender Identity and Transgender Identity on Identity as a Man (Multi-Component Social Identity Scale, Leach et al. 2008) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	4.27	0.209	385	3.8585	4.68
Woman	Yes	2.43	0.209	385	2.0202	2.84
Nonbin	Yes	3.00	0.115	385	2.7710	3.22
Man	No	5.03	0.129	385	4.7803	5.29
Woman	No	4.28	0.139	385	4.0053	4.55
Nonbin	No	2.57	1.274	385	0.0674	5.08
Man	DK/NA	4.32	0.481	385	3.3745	5.27
Woman	DK/NA	2.58	0.570	385	1.4552	3.69
Nonbin	DK/NA	3.15	0.735	385	1.7025	4.59

Note. Summgen3: Condensed multiple-choice gender identity (Man = Man, Woman = Woman, Nonbin = Non-binary or Agender). Transid3: Condensed transgender identity (Yes = Transgender, No = Cisgender, DK/NA = Don't know or Choose not to answer).

Table C4

*Fixed-Effects ANOVA for Gender Identity on Identity as a Man (Multi-Dimensional Social Identity Scale, Leach et al. 2008) - Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	3246.67	1	3246.67	1727.68	.000		
summgen3	225.67	2	112.83	60.04	.000	.23	[.17, .29]
Error	736.65	392	1.88				

Note. LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C5

*Fixed-Effects ANOVA for Transgender Identity on Identity as a Man (Multi-Dimensional Social Identity Scale, Leach et al. 2008) - Study 1.*



Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	4027.18	1	4027.18	2176.71	.000		
transid3	246.64	2	123.32	66.65	.000	.23	[.18, .29]
Error	808.50	437	1.85				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C6

*Fixed-Effects ANOVAs for Interactions of Gender Identity and Transgender Identity on Identity as a Woman (Multi-Dimensional Social Identity Scale, Leach et al. 2008) - Study 1.*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1786.07	1	1786.07	1321.33	.000		
summgen3	66.23	2	33.12	24.50	.000	.11	[.06, .16]
transid3	28.27	2	14.13	10.46	.000	.05	[.02, .09]
summgen3 x transid3	16.58	4	4.14	3.07	.017	.03	[.00, .05]
Error	528.52	391	1.35				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C7

*Expected Means for Gender Identity and Transgender Identity on Identity as a Woman (Multi-Component Social Identity Scale, Leach et al. 2008) – Study 1.*

summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
Man	Yes	3.40	0.194	391	3.02	3.78
Woman	Yes	5.33	0.184	391	4.97	5.69
Nonbin	Yes	3.96	0.104	391	3.76	4.17
Man	No	4.43	0.122	391	4.19	4.67
Woman	No	5.62	0.121	391	5.39	5.86
Nonbin	No	4.00	1.163	391	1.71	6.29
Man	DK/NA	4.52	0.439	391	3.66	5.39
Woman	DK/NA	4.34	0.475	391	3.40	5.27
Nonbin	DK/NA	4.74	0.671	391	3.42	6.06

*Note.* Summgen3: Condensed multiple-choice gender identity (Man = Man, Woman = Woman, Nonbin = Non-binary or Agender). Transid3: Condensed transgender identity (Yes = Transgender, No = Cisgender, DK/NA = Don't know or Choose not to answer).

Table C8

*Fixed-Effects ANOVA for Gender Identity on Identity as a Woman (Multi-Dimensional Social Identity Scale, Leach et al. 2008) - Study 1.*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2317.66	1	2317.66	1618.82	.000		
summgen3	183.48	2	91.74	64.08	.000	.24	[.18, .30]
Error	569.82	398	1.43				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C9

*Fixed-Effects ANOVA for Transgender Identity on Identity as a Woman (Multi-Dimensional Social Identity Scale, Leach et al. 2008) - Study 1.*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	4684.26	1	4684.26	2853.06	.000		
transid3	91.58	2	45.79	27.89	.000	.11	[.07, .16]
Error	727.33	443	1.64				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C10

*Expected Means for Gender Identity and Transgender Identity on Identification with Men (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	5.77	0.290	216	5.200	6.34
Woman	Yes	2.00	0.275	216	1.459	2.54
Nonbin	Yes	3.55	0.153	216	3.250	3.85
Man	No	6.44	0.572	216	5.318	7.57
Woman	No	3.50	0.606	216	2.305	4.70
Nonbin	No	5.00	1.715	216	1.620	8.38
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	1.00	0.990	216	-0.952	2.95
Nonbin	DK/NA	4.00	0.858	216	2.310	5.69

*Note.* Summgen3: Condensed multiple-choice gender identity (Man = Man, Woman = Woman, Nonbin = Non-binary or Agender). Transid3: Condensed transgender identity (Yes = Transgender, No = Cisgender, DK/NA = Don't know or Choose not to answer).

Table C11

*Fixed-Effects ANOVA for Gender Identity on Identification with Men (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1536.36	1	1536.36	514.27	.000		
summgen3	332.58	2	166.29	55.66	.000	.33	[.25, .40]
Error	663.21	222	2.99				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C12

*Fixed-Effects ANOVA for Transgender Identity on Identification with Men (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	441.80	1	441.80	105.61	.000		
transid3	24.19	2	12.10	2.89	.057	.02	[.00, .05]
Error	1112.81	266	4.18				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C13

*Expected Means for Gender Identity and Transgender Identity on Identification with Women (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	3.37	0.256	218	2.87	3.88
Woman	Yes	6.28	0.240	218	5.80	6.75
Nonbin	Yes	4.24	0.135	218	3.97	4.50
Man	No	3.56	0.505	218	2.56	4.55
Woman	No	6.50	0.536	218	5.44	7.56
Nonbin	No	5.00	1.515	218	2.01	7.99
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	5.67	0.875	218	3.94	7.39
Nonbin	DK/NA	5.25	0.758	218	3.76	6.74

*Note.* Summgen3: Condensed multiple-choice gender identity (Man = Man, Woman = Woman, Nonbin = Non-binary or Agender). Transid3: Condensed transgender identity (Yes = Transgender, No = Cisgender, DK/NA = Don't know or Choose not to answer).

Table C14

*Fixed-Effects ANOVA for Gender Identity on Identification with Women (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	511.36	1	511.36	223.71	.000		
summgen3	220.23	2	110.11	48.17	.000	.30	[.22, .37]
Error	512.04	224	2.29				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C15

*Fixed-Effects ANOVA for Transgender Identity on Identification with Women (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	441.80	1	441.80	134.52	.000		
transid3	2.69	2	1.34	0.41	.664	.00	[.00, .02]
Error	880.17	268	3.28				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C16

*Expected Means for Gender Identity and Transgender Identity on Identification with Non-Binary People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	5.00	0.226	217	4.55	5.45
Woman	Yes	4.74	0.214	217	4.32	5.17
Nonbin	Yes	6.61	0.119	217	6.38	6.85
Man	No	2.89	0.446	217	2.01	3.77
Woman	No	4.38	0.473	217	3.44	5.31
Nonbin	No	6.00	1.337	217	3.37	8.63
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	4.67	0.772	217	3.15	6.19
Nonbin	DK/NA	5.50	0.668	217	4.18	6.82

*Note.* Summgen3: Condensed multiple-choice gender identity (Man = Man, Woman = Woman, Nonbin = Non-binary or Agender). Transid3: Condensed transgender identity (Yes = Transgender, No = Cisgender, DK/NA = Don't know or Choose not to answer).

Table C17

*Fixed-Effects ANOVA for Gender Identity on Identification with Non-Binary People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	918.20	1	918.20	480.75	.000		
summgen3	208.65	2	104.33	54.62	.000	.33	[.24, .40]
Error	425.92	223	1.91				

Table C18

*Fixed-Effects ANOVA for Transgender Identity on Identification with Non-Binary People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	259.20	1	259.20	114.38	.000		
transid3	128.15	2	64.08	28.28	.000	.17	[.11, .24]
Error	605.05	267	2.27				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C19

*Expected Means for Gender Identity and Transgender Identity on Identification with Gender Non-Conforming People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	5.37	0.216	218	4.95	5.80
Woman	Yes	5.33	0.205	218	4.93	5.74
Nonbin	Yes	6.50	0.113	218	6.27	6.72
Man	No	3.00	0.426	218	2.16	3.84
Woman	No	4.88	0.452	218	3.98	5.77
Nonbin	No	6.00	1.278	218	3.48	8.52
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	5.33	0.738	218	3.88	6.79
Nonbin	DK/NA	6.00	0.639	218	4.74	7.26

Table C20

*Fixed-Effects ANOVA for Gender Identity on Identification with Gender Non-Conforming People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1050.57	1	1050.57	589.48	.000		
summgen3	109.46	2	54.73	30.71	.000	.22	[.14, .29]
Error	399.21	224	1.78				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C21

*Fixed-Effects ANOVA for Transgender Identity on Identification with Gender Non-Conforming People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	281.25	1	281.25	154.15	.000		
transid3	107.71	2	53.85	29.52	.000	.18	[.11, .24]
Error	488.97	268	1.82				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C22

*Expected Means for Gender Identity and Transgender Identity on Identification with Gender Minorities (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	6.03	0.203	218	5.63	6.43
Woman	Yes	5.87	0.192	218	5.49	6.25
Nonbin	Yes	6.28	0.106	218	6.07	6.49
Man	No	3.00	0.400	218	2.21	3.79
Woman	No	5.00	0.424	218	4.16	5.84
Nonbin	No	6.00	1.200	218	3.64	8.36
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	4.00	0.693	218	2.64	5.36
Nonbin	DK/NA	5.50	0.600	218	4.32	6.68

Table C23

*Fixed-Effects ANOVA for Gender Identity on Identification with Gender Minorities (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1287.36	1	1287.36	728.71	.000		
summgen3	30.74	2	15.37	8.70	.000	.07	[.02, .13]
Error	395.72	224	1.77				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C24

*Fixed-Effects ANOVA for Transgender Identity on Identification with Gender Minorities (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	304.20	1	304.20	204.55	.000		
transid3	111.82	2	55.91	37.59	.000	.22	[.15, .28]
Error	398.56	268	1.49				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.



Table C25

*Expected Means for Gender Identity and Transgender Identity on Identification with Transgender People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	6.37	0.269	217	5.842	6.90
Woman	Yes	5.68	0.258	217	5.176	6.19
Nonbin	Yes	5.46	0.141	217	5.187	5.74
Man	No	3.22	0.530	217	2.178	4.27
Woman	No	3.75	0.562	217	2.642	4.86
Nonbin	No	6.00	1.590	217	2.867	9.13
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	2.67	0.918	217	0.858	4.48
Nonbin	DK/NA	4.75	0.795	217	3.183	6.32

Table C26

*Fixed-Effects ANOVA for Gender Identity on Identification with Transgender People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1443.27	1	1443.27	481.25	.000		
summgen3	6.85	2	3.42	1.14	.321	.01	[.00, .04]
Error	668.78	223	3.00				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C27

*Fixed-Effects ANOVA for Transgender Identity on Identification with Transgender People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	231.20	1	231.20	93.25	.000		
transid3	138.45	2	69.22	27.92	.000	.17	[.11, .24]
Error	661.98	267	2.48				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C28

*Expected Means for Gender Identity and Transgender Identity on Identification with Cisgender People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	2.63	0.278	217	2.080	3.18
Woman	Yes	3.03	0.264	217	2.506	3.55
Nonbin	Yes	2.40	0.147	217	2.116	2.69
Man	No	4.67	0.549	217	3.584	5.75
Woman	No	6.00	0.582	217	4.852	7.15
Nonbin	No	3.00	1.647	217	-0.247	6.25
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	2.33	0.951	217	0.459	4.21
Nonbin	DK/NA	3.75	0.824	217	2.127	5.37

Table C29

*Fixed-Effects ANOVA for Gender Identity on Identification with Cisgender People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	408.09	1	408.09	132.09	.000		
summgen3	40.85	2	20.43	6.61	.002	.06	[.01, .11]
Error	688.96	223	3.09				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C30

*Fixed-Effects ANOVA for Transgender Identity on Identification with Cisgender People (Single-Item Social Identification Scales, Postmes et al. 2013) – Study 1.*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	451.25	1	451.25	153.94	.000		
transid3	92.18	2	46.09	15.72	.000	.11	[.05, .16]
Error	782.65	267	2.93				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C31

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Identity as a Man (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	772.60	1	772.60	1723.62	.000		
summgen3	60.95	2	30.48	67.99	.000	.20	[.15, .25]
transid3	1.38	2	0.69	1.54	.215	.01	[.00, .02]
summgen3 x transid3	6.22	4	1.55	3.47	.008	.02	[.00, .04]
Error	243.84	544	0.45				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C32

*Expected Means for Gender Identity and Transgender Identity on Identity as a Man (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	4.03	0.2117	544	3.61	4.44
2	Woman	Yes	1.68	0.2117	544	1.26	2.09
3	Nonbin/Agender	Yes	2.45	0.1165	544	2.22	2.68
4	Man	No	3.92	0.0454	544	3.83	4.01
5	Woman	No	1.84	0.0422	544	1.76	1.92
6	Nonbin/Agender	No	2.78	0.2232	544	2.34	3.22
7	Man	DK/NA	3.14	0.2232	544	2.70	3.58
8	Woman	DK/NA	2.29	0.2733	544	1.75	2.83
9	Nonbin/Agender	DK/NA	2.29	0.2531	544	1.79	2.78

Table C33

*Fixed-Effects ANOVA for Gender Identity on Identity as a Man (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2463.95	1	2463.95	5380.38	.000		
summgen3	533.09	2	266.55	582.04	.000	.68	[.64, .70]
Error	253.71	554	0.46				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C34

*Fixed-Effects ANOVA for Transgender Identity on Identity as a Man (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1102.82	1	1102.82	786.25	.000		
transid3	1.61	2	0.80	0.57	.563	.00	[.00, .01]
Error	789.69	563	1.40				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C35

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Identity as a Woman (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	843.59	1	843.59	1619.08	.000		
summgen3	63.92	2	31.96	61.34	.000	.19	[.14, .23]
transid3	0.52	2	0.26	0.50	.606	.00	[.00, .01]
summgen3 x transid3	5.96	4	1.49	2.86	.023	.02	[.00, .04]
Error	281.36	540	0.52				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C36

*Expected Means for Gender Identity and Transgender Identity on Identity as a Woman (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	1.82	0.2283	540	1.38	2.27
2	Woman	Yes	4.27	0.2283	540	3.83	4.72
3	Nonbin/Agender	Yes	2.62	0.1257	540	2.37	2.87
4	Man	No	1.84	0.0495	540	1.74	1.94
5	Woman	No	3.98	0.0455	540	3.89	4.07
6	Nonbin/Agender	No	2.47	0.2406	540	2.00	2.94
7	Man	DK/NA	2.33	0.2406	540	1.86	2.81
8	Woman	DK/NA	3.29	0.2947	540	2.71	3.87
9	Nonbin/Agender	DK/NA	2.86	0.2728	540	2.32	3.39

Table C37

*Fixed-Effects ANOVA for Gender Identity on Identity as a Woman (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2618.07	1	2618.07	4964.03	.000		
summgen3	562.58	2	281.29	533.34	.000	.66	[.62, .69]
Error	290.08	550	0.53				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C38

*Fixed-Effects ANOVA for Transgender Identity on Identity as a Woman (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1187.05	1	1187.05	770.88	.000		
transid3	4.34	2	2.17	1.41	.245	.01	[.00, .02]
Error	860.78	559	1.54				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C39

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Identity as a Non-Binary Person (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	452.38	1	452.38	1097.15	.000		
summgen3	51.13	2	25.57	62.01	.000	.19	[.14, .23]
transid3	9.75	2	4.88	11.82	.000	.04	[.02, .07]
summgen3 x transid3	1.70	4	0.42	1.03	.392	.01	[.00, .02]
Error	222.65	540	0.41				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C40

*Expected Means for Gender Identity and Transgender Identity on Identity as a Non-Binary Person (Study 2).*

	summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
1	Man	Yes	2.03	0.2031	540	1.63	2.43
2	Woman	Yes	1.57	0.2031	540	1.17	1.97
3	Nonbin/Agender	Yes	3.48	0.1101	540	3.26	3.70
4	Man	No	1.26	0.0439	540	1.17	1.34
5	Woman	No	1.30	0.0405	540	1.22	1.38
6	Nonbin/Agender	No	2.74	0.2140	540	2.32	3.16
7	Man	DK/NA	1.75	0.2270	540	1.30	2.20
8	Woman	DK/NA	1.61	0.2621	540	1.10	2.13
9	Nonbin/Agender	DK/NA	3.10	0.2427	540	2.62	3.57

Table C41

*Fixed-Effects ANOVA for Gender Identity on Identity as a Non-Binary Person (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1296.62	1	1296.62	2979.86	.000		
summgen3	190.04	2	95.02	218.37	.000	.44	[.39, .48]
Error	239.32	550	0.44				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C42

*Fixed-Effects ANOVA for Transgender Identity on Identity as a Non-Binary Person (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	652.61	1	652.61	1222.39	.000		
transid3	126.95	2	63.48	118.89	.000	.30	[.25, .34]
Error	298.44	559	0.53				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.



Table C43

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Frequency of Being Upset When Others Use Male Pronouns (He/Him) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	221.64	1	221.64	198.39	.000		
summgen3	12.65	2	6.33	5.66	.004	.04	[.01, .07]
transid3	10.75	2	5.38	4.81	.009	.03	[.00, .06]
summgen3 x transid3	13.32	4	3.33	2.98	.019	.04	[.00, .07]
Error	340.74	305	1.12				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C44

*Expected Means for Gender Identity and Transgender Identity on Frequency of Being Upset When Others Use Male Pronouns (He/Him) for You (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	1.44	0.3523	305	0.751	2.14
2	Woman	Yes	3.60	0.3342	305	2.942	4.26
3	Nonbin/Agender	Yes	2.09	0.2253	305	1.647	2.53
4	Man	No	1.22	0.0788	305	1.067	1.38
5	Woman	No	1.92	0.1205	305	1.685	2.16
6	Nonbin/Agender	No	1.80	0.4727	305	0.870	2.73
7	Man	DK/NA	1.86	0.3995	305	1.071	2.64
8	Woman	DK/NA	2.00	0.7474	305	0.529	3.47
9	Nonbin/Agender	DK/NA	3.00	0.7474	305	1.529	4.47

Table C45

*Fixed-Effects ANOVA for Gender Identity on Frequency of Being Upset When Others Use Male Pronouns (He/Him) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	651.26	1	651.26	535.16	.000		
summgen3	60.38	2	30.19	24.81	.000	.14	[.08, .19]
Error	382.12	314	1.22				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C46

*Fixed-Effects ANOVA for Transgender Identity on Frequency of Being Upset When Others Use Male Pronouns (He/Him) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	325.97	1	325.97	259.20	.000		
transid3	31.38	2	15.69	12.48	.000	.07	[.03, .12]
Error	401.18	319	1.26				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C47

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Frequency of Being Upset When Others Use Female Pronouns (She/Her) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	401.78	1	401.78	466.56	.000		
summgen3	42.85	2	21.43	24.88	.000	.12	[.07, .17]
transid3	12.45	2	6.22	7.23	.001	.04	[.01, .07]
summgen3 x transid3	17.51	4	4.38	5.08	.001	.05	[.01, .08]
Error	322.07	374	0.86				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C48

*Expected Means for Gender Identity and Transgender Identity on Frequency of Being Upset When Others Use Female Pronouns (She/Her) for You (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	3.70	0.2935	374	3.123	4.28
2	Woman	Yes	1.00	0.2935	374	0.423	1.58
3	Nonbin/Agender	Yes	3.45	0.1615	374	3.137	3.77
4	Man	No	1.99	0.0984	374	1.795	2.18
5	Woman	No	1.13	0.0626	374	1.004	1.25
6	Nonbin/Agender	No	2.83	0.3788	374	2.088	3.58
7	Man	DK/NA	2.40	0.4150	374	1.584	3.22
8	Woman	DK/NA	1.50	0.4640	374	0.588	2.41
9	Nonbin/Agender	DK/NA	2.33	0.3788	374	1.588	3.08

Table C49

*Fixed-Effects ANOVA for Gender Identity on Frequency of Being Upset When Others Use Female Pronouns (She/Her) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1219.83	1	1219.83	1253.12	.000		
summgen3	210.04	2	105.02	107.89	.000	.36	[.30, .41]
Error	372.83	383	0.97				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C50

*Fixed-Effects ANOVA for Transgender Identity on Frequency of Being Upset When Others Use Female Pronouns (She/Her) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	580.42	1	580.42	504.56	.000		
transid3	127.29	2	63.65	55.33	.000	.22	[.16, .27]
Error	450.93	392	1.15				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C51

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Frequency of Being Upset When Others Use Gender Neutral Pronouns (They/Them) for You (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	149.94	1	149.94	177.63	.000		
summgen3	8.37	2	4.18	4.96	.008	.03	[.00, .07]
transid3	0.73	2	0.36	0.43	.651	.00	[.00, .02]
summgen3 x transid3	2.47	4	0.62	0.73	.571	.01	[.00, .02]
Error	261.67	310	0.84				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C52

*Expected Means for Gender Identity and Transgender Identity on Frequency of Being Upset When Others Use Gender Neutral Pronouns (They/Them) for You (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	2.13	0.3248	310	1.486	2.76
2	Woman	Yes	1.22	0.3063	310	0.620	1.82
3	Nonbin/Agender	Yes	1.10	0.1650	310	0.772	1.42
4	Man	No	1.54	0.0842	310	1.372	1.70
5	Woman	No	1.39	0.0809	310	1.228	1.55
6	Nonbin/Agender	No	1.00	0.3248	310	0.361	1.64
7	Man	DK/NA	1.83	0.3751	310	1.095	2.57
8	Woman	DK/NA	1.33	0.5304	310	0.290	2.38
9	Nonbin/Agender	DK/NA	1.17	0.3751	310	0.429	1.90

Table C53

*Fixed-Effects ANOVA for Gender Identity on Frequency of Being Upset When Others Use Gender Neutral Pronouns (They/Them) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	459.58	1	459.58	549.13	.000		
summgen3	9.73	2	4.87	5.82	.003	.04	[.01, .07]
Error	266.98	319	0.84				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C54

*Fixed-Effects ANOVA for Transgender Identity on Frequency of Being Upset When Others Use Gender Neutral Pronouns (They/Them) for You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	206.35	1	206.35	243.81	.000		
transid3	1.05	2	0.52	0.62	.537	.00	[.00, .02]
Error	274.21	324	0.85				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C55

*Expected Means for Gender Identity and Transgender Identity on Frequency of Discomfort When Required to Check the “M” Box (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	2.38	0.2789	299	1.826	2.92
2	Woman	Yes	4.00	0.3221	299	3.366	4.63
3	Nonbin/Agender	Yes	2.00	0.2630	299	1.483	2.52
4	Man	No	1.23	0.0563	299	1.119	1.34
5	Woman	No	1.36	0.0911	299	1.181	1.54
6	Nonbin/Agender	No	3.25	0.3944	299	2.474	4.03
7	Man	DK/NA	1.75	0.2789	299	1.201	2.30
8	Woman	DK/NA	2.00	0.7889	299	0.448	3.55
9	Nonbin/Agender	DK/NA	NA	NA	NA	NA	NA

Table C56

*Fixed-Effects ANOVA for Gender Identity on Frequency of Discomfort When Required to Check the “M” Box (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	364.33	1	364.33	447.92	.000		
summgen3	28.23	2	14.12	17.35	.000	.10	[.05, .15]
Error	249.71	307	0.81				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C57

*Fixed-Effects ANOVA for Transgender Identity on Frequency of Discomfort When Required to Check the “M” Box (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	255.92	1	255.92	337.34	.000		
transid3	51.98	2	25.99	34.26	.000	.18	[.12, .24]
Error	235.18	310	0.76				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C58

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Frequency of Discomfort When Required to Check the “F” Box (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	546.56	1	546.56	645.94	.000		
summgen3	85.45	2	42.73	50.49	.000	.22	[.16, .28]
transid3	17.67	2	8.84	10.44	.000	.05	[.02, .09]
summgen3 x transid3	28.08	4	7.02	8.30	.000	.08	[.04, .12]
Error	303.77	359	0.85				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C59

*Expected Means for Gender Identity and Transgender Identity on Frequency of Discomfort When Required to Check the “F” Box (Study 2).*

	summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
1	Man	Yes	4.12	0.3252	359	3.485	4.76
2	Woman	Yes	1.80	0.4114	359	0.991	2.61
3	Nonbin/Agender	Yes	4.10	0.1652	359	3.772	4.42
4	Man	No	1.51	0.1116	359	1.295	1.73
5	Woman	No	1.33	0.0604	359	1.209	1.45
6	Nonbin/Agender	No	4.17	0.3755	359	3.428	4.91
7	Man	DK/NA	1.86	0.3477	359	1.173	2.54
8	Woman	DK/NA	1.25	0.4599	359	0.345	2.15
9	Nonbin/Agender	DK/NA	4.14	0.3477	359	3.459	4.83



Table C60

*Fixed-Effects ANOVA for Gender Identity on Frequency of Discomfort When Required to Check the “F” Box (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1410.09	1	1410.09	1426.43	.000		
sumngen3	302.29	2	151.15	152.90	.000	.45	[.39, .50]
Error	364.77	369	0.99				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C61

*Fixed-Effects ANOVA for Transgender Identity on Frequency of Discomfort When Required to Check the “F” Box (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	856.02	1	856.02	731.64	.000		
transid3	235.37	2	117.69	100.59	.000	.35	[.29, .40]
Error	438.75	375	1.17				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C62

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Frequency of Discomfort When Any Gender Information is Requested (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	699.28	1	699.28	778.90	.000		
summgen3	51.94	2	25.97	28.93	.000	.11	[.06, .15]
transid3	58.90	2	29.45	32.80	.000	.12	[.08, .16]
summgen3 x transid3	19.89	4	4.97	5.54	.000	.04	[.01, .07]
Error	437.22	487	0.90				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C63

*Expected Means for Gender Identity and Transgender Identity on Frequency of Discomfort When Any Gender Information is Requested (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	3.45	0.2857	487	2.893	4.02
2	Woman	Yes	3.75	0.3350	487	3.092	4.41
3	Nonbin/Agender	Yes	3.97	0.1625	487	3.651	4.29
4	Man	No	1.41	0.0706	487	1.267	1.54
5	Woman	No	1.53	0.0618	487	1.406	1.65
6	Nonbin/Agender	No	3.67	0.3158	487	3.046	4.29
7	Man	DK/NA	1.62	0.3350	487	0.967	2.28
8	Woman	DK/NA	1.75	0.4738	487	0.819	2.68
9	Nonbin/Agender	DK/NA	3.71	0.3581	487	3.011	4.42

Table C64

*Fixed-Effects ANOVA for Gender Identity on Frequency of Discomfort When Any Gender Information is Requested (Study 2)*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1734.30	1	1734.30	1634.67	.000		
summgen3	245.66	2	122.83	115.77	.000	.32	[.26, .37]
Error	527.29	497	1.06				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C65

*Fixed-Effects ANOVA for Transgender Identity on Frequency of Discomfort When Any Gender Information is Requested (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	994.65	1	994.65	938.73	.000		
transid3	276.36	2	138.18	130.41	.000	.34	[.29, .39]
Error	535.08	505	1.06				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C66

*Fixed-Effects ANOVAs for Interactions of Gender Identity and Transgender Identity on Positive Emotions Toward Own Gender Group (Study 1).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1492.58	1	1492.58	797.76	.000		
summgen3	52.40	2	26.20	14.00	.000	.07	[.03, .10]
transid3	12.85	2	6.42	3.43	.033	.02	[.00, .04]
summgen3 x transid3	2.60	4	0.65	0.35	.845	.00	[.00, .01]
Error	750.25	401	1.87				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C67

*Expected Means for Gender Identity and Transgender Identity on Positive Emotions toward Other People with the Same Gender as You (Study 1).*

summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
Man	Yes	4.53	0.225	401	4.09	4.97
woman	Yes	5.65	0.222	401	5.22	6.09
Nonbin	Yes	5.80	0.121	401	5.57	6.04
Man	No	3.92	0.139	401	3.65	4.20
woman	No	4.97	0.142	401	4.69	5.25
Nonbin	No	4.00	1.368	401	1.31	6.69
Man	DK/NA	3.43	0.517	401	2.41	4.44
woman	DK/NA	5.12	0.558	401	4.03	6.22
Nonbin	DK/NA	4.88	0.684	401	3.53	6.22

Table C68

*Fixed-Effects ANOVA for Gender Identity on Positive Emotions Toward Own Gender Group (Study 1).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2321.81	1	2321.81	1208.94	.000		
sumngen3	207.93	2	103.97	54.13	.000	.21	[.15, .26]
Error	783.58	408	1.92				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C69

*Fixed-Effects ANOVA for Transgender Identity on Positive Emotions Toward Own Gender Group (Study 1).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	3782.14	1	3782.14	1790.39	.000		
transid3	133.32	2	66.66	31.56	.000	.12	[.08, .17]
Error	956.95	453	2.11				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C70

*Fixed-Effects ANOVAs for Interactions of Gender Identity and Transgender Identity on Negative Emotions Toward Own Gender Group (Study 1).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	610.93	1	610.93	341.46	.000		
summgen3	2.68	2	1.34	0.75	.474	.00	[.00, .02]
transid3	2.78	2	1.39	0.78	.461	.00	[.00, .02]
summgen3 x transid3	7.37	4	1.84	1.03	.391	.01	[.00, .02]
Error	715.67	400	1.79				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C71

*Expected Means for Gender Identity and Transgender Identity on Negative Emotions toward Other People with the Same Gender as You (Study 1).*

summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
Man	Yes	2.75	0.220	400	2.321	3.19
woman	Yes	2.28	0.217	400	1.854	2.71
Nonbin	Yes	2.09	0.119	400	1.854	2.32
Man	No	2.51	0.136	400	2.243	2.78
woman	No	2.42	0.139	400	2.152	2.70
Nonbin	No	4.00	1.338	400	1.370	6.63
Man	DK/NA	3.00	0.506	400	2.006	3.99
woman	DK/NA	3.31	0.546	400	2.232	4.38
Nonbin	DK/NA	2.21	0.669	400	0.894	3.52

Table C72

*Fixed-Effects ANOVA for Gender Identity on Negative Emotions Toward Own Gender Group (Study 1).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	951.60	1	951.60	531.44	.000		
summgen3	17.41	2	8.71	4.86	.008	.02	[.00, .05]
Error	728.78	407	1.79				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C73

*Fixed-Effects ANOVA for Transgender Identity on Negative Emotions Toward Own Gender Group (Study 1).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1169.23	1	1169.23	640.31	.000		
transid3	15.23	2	7.62	4.17	.016	.02	[.00, .04]
Error	825.37	452	1.83				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C74

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Positive Emotions toward Other People with the Same Gender as You (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2462.09	1	2462.09	859.01	.000		
summgen3	41.78	2	20.89	7.29	.001	.03	[.01, .05]
transid3	31.86	2	15.93	5.56	.004	.02	[.00, .04]
summgen3 x transid3	28.86	4	7.21	2.52	.040	.02	[.00, .03]
Error	1556.35	543	2.87				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C75

*Expected Means for Gender Identity and Transgender Identity on Positive Emotions toward Other People with the Same Gender as You (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	5.40	0.535	543	4.35	6.45
2	Woman	Yes	4.73	0.535	543	3.67	5.78
3	Nonbin/Agender	Yes	5.99	0.290	543	5.42	6.56
4	Man	No	3.40	0.115	543	3.17	3.62
5	Woman	No	4.55	0.107	543	4.34	4.76
6	Nonbin/Agender	No	4.97	0.564	543	3.86	6.08
7	Man	DK/NA	4.50	0.564	543	3.39	5.61
8	Woman	DK/NA	4.08	0.691	543	2.73	5.44
9	Nonbin/Agender	DK/NA	6.58	0.691	543	5.23	7.94



Table C76

*Fixed-Effects ANOVA for Gender Identity on Positive Emotions toward Other People with the Same Gender as You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	7063.86	1	7063.86	2401.17	.000		
summgen3	276.84	2	138.42	47.05	.000	.15	[.10, .19]
Error	1626.84	553	2.94				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C77

*Fixed-Effects ANOVA for Transgender Identity on Positive Emotions toward Other People with the Same Gender as You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	3510.15	1	3510.15	1105.91	.000		
transid3	137.03	2	68.52	21.59	.000	.07	[.04, .11]
Error	1783.78	562	3.17				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C78

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Negative Emotions toward Other People with the Same Gender as You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	493.45	1	493.45	241.41	.000		
summgen3	30.69	2	15.35	7.51	.001	.03	[.01, .05]
transid3	7.07	2	3.54	1.73	.178	.01	[.00, .02]
summgen3 x transid3	10.11	4	2.53	1.24	.294	.01	[.00, .02]
Error	1097.63	537	2.04				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C79

*Expected Means for Gender Identity and Transgender Identity on Negative Emotions toward Other People with the Same Gender as You (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	2.60	0.4521	537	1.712	3.49
2	Woman	Yes	1.85	0.4521	537	0.962	2.74
3	Nonbin/Agender	Yes	1.77	0.2452	537	1.288	2.25
4	Man	No	2.33	0.0982	537	2.136	2.52
5	Woman	No	2.22	0.0902	537	2.045	2.40
6	Nonbin/Agender	No	1.31	0.5055	537	0.320	2.31
7	Man	DK/NA	3.63	0.4766	537	2.693	4.57
8	Woman	DK/NA	2.83	0.5837	537	1.687	3.98
9	Nonbin/Agender	DK/NA	1.42	0.5837	537	0.270	2.56

Table C80

*Fixed-Effects ANOVA for Gender Identity on Negative Emotions toward Other People with the Same Gender as You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1409.09	1	1409.09	686.94	.000		
summgen3	22.17	2	11.09	5.40	.005	.02	[.00, .04]
Error	1122.04	547	2.05				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C81

*Fixed-Effects ANOVA for Transgender Identity on Negative Emotions toward Other People with the Same Gender as You (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	810.98	1	810.98	387.89	.000		
transid3	14.20	2	7.10	3.40	.034	.01	[.00, .03]
Error	1162.46	556	2.09				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C82

*Expected Means for Gender Identity and Transgender Identity on Gender Centrality (Study 1).*

sumngen3	transid3	emmean	SE	df	lower.CL	upper.CL
Man	Yes	5.94	0.197	218	5.55	6.33
Woman	Yes	6.19	0.187	218	5.82	6.56
Nonbin	Yes	5.97	0.104	218	5.76	6.17
Man	No	4.44	0.389	218	3.68	5.21
Woman	No	5.46	0.413	218	4.64	6.27
Nonbin	No	4.67	1.167	218	2.37	6.97
Man	DK/NA	nonEst	NA	NA	NA	NA
Woman	DK/NA	5.33	0.674	218	4.01	6.66
Nonbin	DK/NA	4.50	0.584	218	3.35	5.65

Table C83

*Fixed-Effects ANOVA for Gender Identity on Gender Centrality (Study 1).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1397.82	1	1397.82	951.14	.000		
sumngen3	3.79	2	1.90	1.29	.277	.01	[.00, .04]
Error	329.20	224	1.47				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C84

*Fixed -Effects ANOVAs for Gender Identity and Transgender Identity on Gender Identification (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1876.60	1	1876.60	1636.73	.000		
summgen3	5.47	2	2.73	2.39	.093	.01	[.00, .02]
transid3	11.28	2	5.64	4.92	.008	.02	[.00, .04]
summgen3 x transid3	21.23	4	5.31	4.63	.001	.03	[.01, .05]
Error	628.31	548	1.15				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C85

*Expected Means for Gender Identity and Transgender Identity on Gender Identification (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	4.00	0.3229	548	3.37	4.63
2	Woman	Yes	4.13	0.3386	548	3.47	4.80
3	Nonbin/Agender	Yes	4.23	0.1864	548	3.86	4.59
4	Man	No	5.31	0.0729	548	5.17	5.45
5	Woman	No	5.17	0.0668	548	5.04	5.30
6	Nonbin/Agender	No	3.50	0.3569	548	2.80	4.20
7	Man	DK/NA	4.13	0.3569	548	3.43	4.83
8	Woman	DK/NA	3.83	0.4371	548	2.97	4.69
9	Nonbin/Agender	DK/NA	4.08	0.4371	548	3.22	4.94

Table C86

*Fixed-Effects ANOVA for Gender Identity on Gender Identification (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	7422.68	1	7422.68	5967.55	.000		
sumngen3	51.18	2	25.59	20.57	.000	.07	[.04, .10]
Error	694.06	558	1.24				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C87

*Fixed-Effects ANOVA for Transgender Identity on Gender Identification (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2913.94	1	2913.94	2477.78	.000		
transid3	82.43	2	41.22	35.04	.000	.11	[.07, .15]
Error	666.81	567	1.18				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C88

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Use of Male Pronouns (He/Him) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	640.26	1	640.26	535.53	.000		
summgen3	94.22	2	47.11	39.40	.000	.13	[.08, .17]
transid3	18.40	2	9.20	7.69	.001	.03	[.01, .05]
summgen3 x transid3	24.57	4	6.14	5.14	.000	.04	[.01, .06]
Error	656.36	549	1.20				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C89

*Expected Means for Gender Identity and Transgender Identity on Use of Male Pronouns (He/Him) to Refer to Yourself (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	4.10	0.3458	549	3.421	4.78
2	Woman	Yes	2.10	0.3458	549	1.421	2.78
3	Nonbin/Agender	Yes	2.15	0.1903	549	1.778	2.53
4	Man	No	4.21	0.0739	549	4.069	4.36
5	Woman	No	1.14	0.0685	549	1.003	1.27
6	Nonbin/Agender	No	3.11	0.3645	549	2.395	3.83
7	Man	DK/NA	2.78	0.3645	549	2.062	3.49
8	Woman	DK/NA	1.33	0.4464	549	0.457	2.21
9	Nonbin/Agender	DK/NA	1.29	0.4133	549	0.474	2.10

Table C90

*Fixed-Effects ANOVA for Gender Identity on Use of Male Pronouns (He/Him) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2054.33	1	2054.33	1643.11	.000		
summgen3	1138.89	2	569.45	455.46	.000	.62	[.58, .65]
Error	698.90	559	1.25				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C91

*Fixed-Effects ANOVA for Transgender Identity on Use of Male Pronouns (He/Him) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	862.70	1	862.70	265.26	.000		
transid3	6.57	2	3.29	1.01	.365	.00	[.00, .01]
Error	1847.30	568	3.25				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.



Table C92

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Use of Female Pronouns (She/Her) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	784.58	1	784.58	693.29	.000		
summgen3	91.87	2	45.94	40.59	.000	.13	[.09, .17]
transid3	0.72	2	0.36	0.32	.728	.00	[.00, .01]
summgen3 x transid3	25.56	4	6.39	5.65	.000	.04	[.01, .06]
Error	621.29	549	1.13				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C93

*Expected Means for Gender Identity and Transgender Identity on Use of Female Pronouns (She/Her) to Refer to Yourself (Study 2).*

	summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
1	Man	Yes	1.90	0.3364	549	1.239	2.56
2	Woman	Yes	3.50	0.3364	549	2.839	4.16
3	Nonbin/Agender	Yes	2.45	0.1852	549	2.091	2.82
4	Man	No	1.25	0.0719	549	1.105	1.39
5	Woman	No	4.47	0.0666	549	4.340	4.60
6	Nonbin/Agender	No	2.56	0.3546	549	1.859	3.25
7	Man	DK/NA	1.56	0.3546	549	0.859	2.25
8	Woman	DK/NA	3.33	0.4343	549	2.480	4.19
9	Nonbin/Agender	DK/NA	3.57	0.4021	549	2.782	4.36

Table C94

*Fixed-Effects ANOVA for Gender Identity on Use of Female Pronouns (She/Her) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2522.34	1	2522.34	2167.69	.000		
summgen3	1248.94	2	624.47	536.67	.000	.66	[.62, .69]
Error	650.46	559	1.16				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C95

*Fixed-Effects ANOVA for Transgender Identity on Use of Female Pronouns (She/Her) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1106.07	1	1106.07	328.14	.000		
transid3	12.88	2	6.44	1.91	.149	.01	[.00, .02]
Error	1914.58	568	3.37				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C96

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Use of Gender Neutral Pronouns (They/Them) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	df	Mean Square	F	p	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	503.90	1	503.90	522.00	.000		
summgen3	68.39	2	34.20	35.42	.000	.11	[.07, .15]
transid3	6.41	2	3.21	3.32	.037	.01	[.00, .03]
summgen3 x transid3	1.10	4	0.28	0.28	.888	.00	[.00, .00]
Error	530.93	550	0.97				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C97

*Expected Means for Gender Identity and Transgender Identity on Use of Gender Neutral Pronouns (They/Them) to Refer to Yourself (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	1.80	0.3107	550	1.190	2.41
2	Woman	Yes	1.80	0.3107	550	1.190	2.41
3	Nonbin/Agender	Yes	3.82	0.1710	550	3.482	4.15
4	Man	No	1.44	0.0664	550	1.313	1.57
5	Woman	No	1.37	0.0614	550	1.247	1.49
6	Nonbin/Agender	No	3.11	0.3275	550	2.468	3.75
7	Man	DK/NA	1.56	0.3275	550	0.912	2.20
8	Woman	DK/NA	1.67	0.4011	550	0.879	2.45
9	Nonbin/Agender	DK/NA	3.14	0.3714	550	2.413	3.87

Table C98

*Fixed-Effects ANOVA for Gender Identity on Use of Gender Neutral Pronouns (They/Them) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1529.51	1	1529.51	1579.69	.000		
summgen3	225.96	2	112.98	116.69	.000	.29	[.24, .34]
Error	542.21	560	0.97				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C99

*Fixed-Effects ANOVA for Gender Identity on Use of Gender Neutral Pronouns (They/Them) to Refer to Yourself (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	752.54	1	752.54	673.96	.000		
transid3	138.08	2	69.04	61.83	.000	.18	[.13, .22]
Error	635.34	569	1.12				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C100

*Fixed-Effects ANOVA for Gender Identity and Transgender Identity on Choosing Men's Clothing (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	871.71	1	871.71	7432.08	.000		
summgen3	0.81	2	0.40	3.45	.032	.01	[.00, .03]
transid3	0.05	2	0.02	0.21	.810	.00	[.00, .01]
summgen3 x transid3	0.13	4	0.03	0.28	.891	.00	[.00, .00]
Error	64.28	548	0.12				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C101

*Expected Means for Gender Identity and Transgender Identity on Choosing Men's Clothing (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	4.45	0.3026	548	3.856	5.04
2	Woman	Yes	2.30	0.3026	548	1.706	2.89
3	Nonbin/Agender	Yes	3.27	0.1666	548	2.945	3.60
4	Man	No	4.49	0.0647	548	4.359	4.61
5	Woman	No	1.75	0.0601	548	1.634	1.87
6	Nonbin/Agender	No	3.67	0.3190	548	3.040	4.29
7	Man	DK/NA	3.67	0.3190	548	3.040	4.29
8	Woman	DK/NA	1.75	0.3907	548	0.983	2.52
9	Nonbin/Agender	DK/NA	2.64	0.3617	548	1.932	3.35

Table C102

*Fixed-Effects ANOVA for Gender Identity on Choosing Men's Clothing (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	3516.16	1	3516.16	3288.68	.000		
summgen3	760.45	2	380.23	355.63	.000	.56	[.52, .60]
Error	597.67	559	1.07				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C103

*Fixed-Effects ANOVA for Transgender Identity on Choosing Men's Clothing (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1588.27	1	1588.27	660.66	.000		
transid3	4.52	2	2.26	0.94	.391	.00	[.00, .01]
Error	1365.51	568	2.40				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C104

*Fixed-Effects ANOVAs for Gender Identity and Transgender Identity on Choosing Women's Clothing (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1034.94	1	1034.94	1235.07	.000		
summgen3	80.30	2	40.15	47.91	.000	.15	[.10, .19]
transid3	1.01	2	0.50	0.60	.547	.00	[.00, .01]
summgen3 x transid3	32.09	4	8.02	9.57	.000	.06	[.03, .10]
Error	461.72	551	0.84				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C105

*Expected Means for Gender Identity and Transgender Identity on Choosing Women's Clothing (Study 2).*

	summgen3	transid3	emmean	SE	df	lower.CL	upper.CL
1	Man	Yes	1.75	0.2895	551	1.18	2.32
2	Woman	Yes	3.40	0.2895	551	2.83	3.97
3	Nonbin/Agender	Yes	3.11	0.1594	551	2.79	3.42
4	Man	No	1.27	0.0619	551	1.15	1.39
5	Woman	No	4.40	0.0571	551	4.29	4.52
6	Nonbin/Agender	No	2.78	0.3051	551	2.18	3.38
7	Man	DK/NA	2.06	0.3051	551	1.46	2.65
8	Woman	DK/NA	3.00	0.3737	551	2.27	3.73
9	Nonbin/Agender	DK/NA	4.00	0.3460	551	3.32	4.68

Table C106

*Fixed-Effects ANOVA for Gender Identity on Choosing Women's Clothing (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	2870.61	1	2870.61	3044.30	.000		
summgen3	1140.30	2	570.15	604.65	.000	.68	[.65, .71]
Error	528.99	561	0.94				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C107

*Fixed-Effects ANOVA for Transgender Identity on Choosing Women's Clothing (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1271.55	1	1271.55	428.27	.000		
transid3	0.63	2	0.32	0.11	.900	.00	[.00, .00]
Error	1692.37	570	2.97				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.



Table C108

*Fixed-Effects ANOVAs for Interactions of Gender Identity and Transgender Identity on Non-Binary Experiences Scale (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	554.62	1	554.62	533.90	.000		
summgen3	17.13	2	8.56	8.25	.000	.05	[.02, .10]
transid3	7.94	2	3.97	3.82	.023	.03	[.00, .06]
summgen3 x transid3	2.78	4	0.69	0.67	.613	.01	[.00, .02]
Error	307.48	296	1.04				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C109

*Expected Means for Gender Identity and Transgender Identity on Non-Binary Experiences Scale (Study 2).*

	summgen3	transid3	emmean	SE	<i>df</i>	lower.CL	upper.CL
1	Man	Yes	2.43	0.5096	296	1.43	3.44
2	Woman	Yes	2.94	0.4161	296	2.12	3.76
3	Nonbin/Agender	Yes	4.25	0.2548	296	3.75	4.75
4	Man	No	2.21	0.0884	296	2.04	2.39
5	Woman	No	2.18	0.0904	296	2.00	2.35
6	Nonbin/Agender	No	4.11	0.5096	296	3.10	5.11
7	Man	DK/NA	3.70	0.3603	296	2.99	4.41
8	Woman	DK/NA	3.45	0.4558	296	2.56	4.35
9	Nonbin/Agender	DK/NA	4.24	0.7207	296	2.82	5.66

Table C110

*Fixed-Effects ANOVA for Gender Identity on Non-Binary Experiences Scale (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	1347.68	1	1347.68	1219.18	.000		
summgen3	83.58	2	41.79	37.81	.000	.20	[.13, .26]
Error	336.04	304	1.11				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

Table C111

*Fixed-Effects ANOVA for Transgender Identity on Non-Binary Experiences Scale (Study 2).*

Predictor	Sum of Squares	<i>df</i>	Mean Square	<i>F</i>	<i>p</i>	partial $\eta^2$	partial $\eta^2$ 90% CI [LL, UL]
(Intercept)	937.35	1	937.35	849.05	.000		
transid3	80.76	2	40.38	36.58	.000	.19	[.13, .25]
Error	340.03	308	1.10				

*Note.* LL and UL represent the lower-limit and upper-limit of the partial  $\eta^2$  confidence interval, respectively.

## Appendix D

### Additional Measures to Validate Novel Gender Essentialism Scale

#### **Homophobia Scale (Wright, Adams, and Bernat, 1999).**

Participants indicated their agreement with each of 25 items measuring negative attitudes toward homosexual people on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include “I avoid gay individuals” and “I feel that you cannot trust a person who is homosexual.” I averaged item scores into a single composite score ( $\alpha = .95$ ,  $M = 2.3$ ,  $SD = 1$ ); higher scores indicate more homophobic attitudes.

#### **Rosenberg Self-Esteem Scale (Rosenberg, 1965).**

Participants indicated their agreement with each of 10 items measuring self-esteem on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include “I certainly feel useless at times (reverse-coded)” and “I take a positive attitude toward myself.” I averaged item scores into a single composite score ( $\alpha = .89$ ,  $M = 4.8$ ,  $SD = 1.1$ ); higher scores indicate higher self-esteem.

#### **Social Desirability Short Form (Strahan & Gerbasi, 1972).**

Participants indicated whether each of 10 items were true or false (1 = true, 2 = false) of them. These items, such as “I’m always willing to admit it when I make a mistake.” and “I never resent being asked to return a favor,” measure the degree to which a participant’s responses are influenced by social desirability. I averaged item scores into a single composite score ( $\alpha = .48$ ,  $M = 1.3$ ,  $SD = .19$ ).

#### **Need for Closure Short Form (Roets & Ban Hiel, 2011).**

Participants indicated their agreement with each of 15 items measuring need for closure on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include

“I don't like situations that are uncertain” and “I dislike questions which could be answered in many different ways.” I averaged item scores into a single composite score ( $\alpha = .89$ ,  $M = 4.4$ ,  $SD = .90$ ); higher scores indicate more need for closure.

**General Essentialism Scale (Bastian & Haslam, 2006).**

Participants indicated their agreement with each of 24 items measuring general essentialist views on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Example items include “The kind of person someone is can be largely attributed to their genetic inheritance” and “Everyone is either a certain type of person or they are not.” I averaged item scores into a single composite score ( $\alpha = .81$ ,  $M = 3.8$ ,  $SD = .56$ ); higher scores indicate more essentialist views.

**Ambivalent Sexism Scale (Glick and Fiske, 1996).**

Participants indicated their agreement with each of 12 items measuring hostile and benevolent sexism on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). Six items measure benevolent sexism (e.g. “Women should be cherished and protected by men.”) and six items measure hostile sexism (e.g. “Women seek to gain power by getting control over men.”). I averaged item scores into a single composite score for hostile sexism ( $\alpha = .89$ ,  $M = 3.2$ ,  $SD = 1.3$ ) and benevolent sexism ( $\alpha = .82$ ,  $M = 3.7$ ,  $SD = 1.2$ ). In each case, higher scores indicate more sexist attitudes.

**Implicit Person Theory Scale (Dweck et al., 1995).**

Participants indicated their agreement with each of three items measuring general entity vs. incremental theories about people on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly disagree). The items are “The kind of person someone is, is something very basic about them and it can't be changed very much,” “People can do things differently, but the important

parts of who they are can't really be changed,” and “Everyone is a certain kind of person and there is not much that can be done to really change that.” I averaged item scores into a single composite score ( $\alpha = .87$ ,  $M = 3.8$ ,  $SD = 1.4$ ); higher scores indicate a greater endorsement of an entity theory of people.

### **Right-Wing Authoritarianism (Altemeyer, 1981).**

Participants indicated their agreement with each of 15 items measuring right-wing authoritarianism on a 7-point Likert scale (1 = Strongly disagree, 7 = Strongly agree).

Example items include “A lot of our rules regarding sexual behavior are just customs which are not necessarily any better or holier than those which other people follow” and “What our country needs is the most is disciplined citizens, following national leaders in unity.” I averaged item scores into a single composite score ( $\alpha = .90$ ,  $M = 4.6$ ,  $SD = .98$ ); higher scores indicate less right-wing authoritarianism.

### **Demographics Questions**

*What is your gender identity?* [single-select]

- *Man*
- *Woman*
- *Non-binary (e.g. Genderqueer)* [free text box]
- *Genderfluid*
- *Agender*
- *Another identity not listed (please specify)* [free text box]
- *Do not know*
- *Choose not to answer*

*What is your sexual orientation?* [single-select]

- *Straight (Heterosexual)*
- *Gay/Lesbian (Homosexual)*
- *Bisexual/Pansexual*
- *Queer* [free text box]
- *Asexual*
- *Another identity not listed (please specify)* [free text box]
- *Do not know*
- *Choose not to answer*

What is your race/ethnicity (please check all that apply)?

- African American/Black
- Pacific Islander (e.g. Hawaii, Guam, Samoa)
- Indian Subcontinent (e.g. India, Pakistan, Sri Lanka, Bangladesh)
- Middle Eastern (e.g. Egypt, Turkey, U.A.E.)
- East Asian (e.g. Japan, China, Korea)
- Southeast Asian (e.g. Indonesia, Thailand, Vietnam, Philippines)
- White (e.g. European, Anglo, Caucasian)
- Hispanic, Latino(a), Chicano(a)
- Native American (e.g. Cherokee, Choctaw, Inuit, Navajo)
- Another race or ethnicity not listed (please specify) [free text box]

What is your age (in years)? [free text box]

Generally speaking, do you usually think of yourself as a Democrat, a Republican, an independent, or what? [single-select]

- Democrat
- Republican
- Independent
- Other party (please specify) [free text box]

Do you consider yourself religious? If so, what is your religion? [single-select]

- No, I do not consider myself religious.
- Yes (please describe) [free text box]

## **Debriefing Materials**

*“Now that data collection is complete, we can give you complete information about the purpose of this study. Broadly, we are studying the effects of differently-worded gender demographic questions on gender essentialism, or the tendency to believe that members of a gender category (e.g. men) share a fixed, underlying “essence” which determines their identity, makes them fundamentally alike, and gives others important information about them. At the beginning of the study, you were randomly assigned to see one of several different types of gender demographic questions. We wanted to determine whether seeing gender demographic questions that allow for multiple selections would reduce participants' endorsement of gender essentialist beliefs relative to questions that only permit participants to choose one option. This effect has been observed for racial demographic questions (Lee et al., 2014), and we hoped to replicate it for gender demographic questions. In addition to the measures of primary interest, we also included measures of several other constructs that have previously been shown to correlate with gender essentialism. We expected these measures to relate to gender essentialism as has been found in previous studies. Thank you for your participation! Please do not share any details about this study with your peers who might take this study until the end of the semester - awareness of the real purpose of our study could invalidate our results.”*

# Margaret Hauptert, PhD

## EDUCATION

- 2019 Ph.D. Psychology, Human Sexuality minor (Kinsey Institute); Indiana University  
**Dissertation title:** Considerations for the development and implementation of transgender-inclusive gender demographic questions
- 2011 B.A. Psychology, Health and Society, Classical Civilization; Beloit College

## PUBLICATIONS

- Mao, J.M., **Hauptert, M.L.**, & Smith, E.R. (2018). How gender identity and transgender status affect perceptions of attractiveness. *Social Psychological and Personality Science*. Advance online publication. doi:10.1177/1948550618783716
- Hauptert, M.L.**, Pope, A.R.D., Garcia, J.R., & Smith, E.R. (2018). An inclusive gender measure. In R. Milhausen, J.K. Sakaluk, T.D. Fisher, C.M. Davis, & W. Yarber (Eds.), *Handbook of Sexuality-Related Measures* (4th ed.). New York: Routledge.
- Hauptert, M.L.**, Moors, A.C., Gesselman, A.N., & Garcia, J.R. (2017). Estimates and correlates of engagement in consensually non-monogamous relationships. *Current Sexual Health Reports*, 9(3), 155-165.
- Hauptert, M.L.**, Gesselman, A.N., Moors, A.C., Fisher, H.E., & Garcia, J.R. (2016). Prevalence of experiences with consensual non-monogamous relationships: Findings from two national samples of single Americans. *Journal of Sex and Marital Therapy*, 43(5), 424-440. doi:10.1080/0092623X.2016.1178675
- Grosfolsky, A., **Hauptert, M.L.**, & Versteeg, S. W. (2011). An exploratory investigation of coffee and lemon scents and olfactory identification. *Perceptual and Motor Skills*, 112(2), 1-3. doi:10.2466/24.PMS.112.2

## AWARDS AND HONORS

- 2017 SSSS Rising Scholar - Society for the Scientific Study of Sexuality  
2017 Outstanding Teaching Assistant, Psychology  
2014, 2015, 2017 Provost's Travel Award for Women in Science  
2014 Clara Mayo Grant – Society for the Psychological Study of Social Issues  
2014 NSF Graduate Research Fellowship Program Honorable Mention

## **CONFERENCE PRESENTATIONS**

- 2017 Forget the Gold Standard: Reframing "Inconsistencies" between Trans-Inclusive Gender Measures (Presentation at SSSS and SPSSI)
- 2016 Prevalence of Experiences with Consensual Non-monogamous Relationships: Findings from Two Nationally Representative Samples of Single Americans (Presentation at SSSS)
- 2016 Wording of Gender Questions Affects Gender Minorities' Investment in Research (Poster at SPSSI)
- 2016 Who I Was, Who I Am and Who I Want To Be: What Can We Learn from Identity Trajectories? (Poster at MPA and SPSP)
- 2015 A Question of Trust: How Context and Wording of Gender Questions Affects Attitudes Toward Researchers (Poster at MPA and SPSP)

## **TEACHING EXPERIENCE**

### **As Instructor of Record**

Introductory Psychology (Spring 2019)  
Social Psychology Capstone (Fall 2018; Spring 2019)

### **As Trainee Instructor**

Introduction to Research Methods (Spring 2016)

### **As Teaching Assistant (Fall 2011 – Spring 2018)**

Introductory Psychology  
Career Planning for Psychology Majors  
Social Psychology  
Developmental Psychology  
Consumer Psychology  
Social Psychology Capstone  
Moral Psychology  
Psychology of Women

## **EXTRACURRICULAR UNIVERSITY SERVICE**

- 2018 SexFest Consensual Non-Monogamy Panel Moderator
- 2016 Presenter for Center for Innovative Teaching and Learning AI Orientation
- 2014 – 2018 Gay, Lesbian, and Bisexual Speakers Bureau Panelist

## **NON-ACADEMIC WORK**

- 2011 – 2013 Technical Communications Training Writer, Epic Systems Corporation

## **COMMUNITY INVOLVEMENT**

- 2013 – 2014 Student Birth Doula, Bloomington Area Birth Services