Subjective Ratings of Gender Dysphoria Scales by Transgender Individuals

M. Paz Galupo[1] · Lex Pulice-Farrow[1]

Abstract

The present research explored transgender individuals' subjective ratings of two clinical measures of gender dysphoria: the Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults (GIDYQ-AA) and the Utrecht Gender Dysphoria Scale (UGDS). Participants read each scale and provided a global rating regarding how well they captured their experiences of gender dysphoria. Participants included 622 transgender individuals who identified as transfeminine ($n = 221$), transmasculine ($n = 206$), and non-binary/agender ($n = 195$). Findings indicated clear patterns of responses across gender identity and assigned sex, but not clinical diagnosis. For the GIDYQ-AA, transfeminine and transmasculine individuals rated the scales more positively than did non-binary/agender individuals. In addition, participants who were assigned male rated the scale to be a more accurate measure of their dysphoria than did participants who were assigned female. For the UGDS, transfeminine individuals rated the scale most positively, followed by transmasculine individuals, and then non-binary/agender individuals. All pairwise comparisons were significant. Likewise, participants who were assigned male rated the scale to be a more accurate measure than did those who were assigned female. It is important to note that subjective ratings were relatively low ($M = 3.40$, SD $= 1.09$ for GIDYQ-AA; $M = 3.43$, SD $= 1.22$ for UGDS on a 5-point scale) where little more than half of the participants (52.5% GIDYQ-AA; 54% UGDS) agreed or strongly agreed that the scales captured their experience. Discussion focused on the implications for using these measures of gender dysphoria in both clinical and research settings.

Keywords Gender dysphoria · Gender identity · Non-binary · Transgender

Introduction

Gender dysphoria has been the central diagnostic lens for gender incongruence in the clinical literature (Drescher, 2015; Drescher, Cohen-Kettenis, & Winter, 2012). Gender dysphoria has been defined as the "distress that may accompany the incongruence between one's experienced or expressed gender and one's assigned gender" (American Psychiatric Association, 2013, p. 451). A diagnosis of gender dysphoria using criteria from either the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; APA, 2013) or the *International Statistical Classification of Diseases and Related Health Programs* (ICD-11; World Health Organization, 2011) is typically determined categorically (i.e., one either does, or does not, meet criterion). Two-dimensional measures of gender dysphoria

are commonly used in clinical settings: the Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults (GIDYQ-AA; Deogracias et al., 2007) and the Utrecht Gender Dysphoria Scale (UGDS; Cohen-Kettenis & van Goozen, 1997). Deogracias et al. argued that a dimensional measure of gender dysphoria has clinical utility for assessing the degree of dysphoria among those experiencing more "subtle signs of gender dysphoria," or who are "subthreshold" (p. 371).

Research Literature Framing Our Understanding of Gender Dysphoria

Current conceptualizations of gender dysphoria have been impacted by the clinical literature that has predominantly focused on diagnosing psychopathology (Drescher, 2010). Early understandings of gender identity relied upon binary conceptualizations of gender/sex[1] where individuals were classified as transsexual based on identifying with the "opposite" of their "genetic" sex (Benjamin, 1966). Early research

✉ M. Paz Galupo
pgalupo@towson.edu

[1] Psychology Department, Towson University, 8000 York
Road, Towson, MD 21252-0001, USA

---

[1] We use gender/sex consistent with van Anders (2015).

focused on developing typologies of transsexuality (Benjamin, 1966; Blanchard, 1989a, 1989b) that centered exclusively on the experiences of trans women.[2]

Based on his experience working with clients in a gender dysphoria program, Fisk (1973, 1974) felt that a focus on typologies was not central to the treatment of gender dysphoria. Instead, he described individuals with dysphoria as being heterogeneous and ranging in severity. He subsequently conceptualized gender dysphoria as a syndrome represented on a continuum with transsexuality as the most extreme form of dysphoria on one end and non-pathological gender on the other (Fisk, 1974). Although clinical diagnoses of gender dysphoria have undoubtedly evolved since the 1970s, this history has framed the way current clinical measures have been conceptualized.

Mainly based on the experiences of trans women, early medical models of transexuality reinforced binary narratives of a "woman trapped in a man's body" (Meyerowitz, 2002) and do not always resonate with the diverse experiences of transgender individuals (Serano, 2010; Veale, Clarke, & Lomax, 2012). These models have been criticized for centering on anatomical classifications and de-emphasizing the role of self-identification (Bockting, Coleman, & Lief, 1991; Devor, 1993). In contrast, psychological research that centers on transgender identity is often framed from a minority stress model (Hendricks & Testa, 2012; Pflum, Testa, Balsam, Goldblum, & Bongar, 2015; Timmins, Rimes, & Rahman, 2017) and this framework has also impacted the way dysphoria is conceptualized.

The minority stress model emphasizes self-identification and an understanding of how stressors from belonging to marginalized social group impacts mental health outcomes (Meyer, 2003, 2015). Research from a minority stress perspective originally situated transgender experience within the larger lesbian, gay, bisexual, and transgender (LGBT) community framework (Fassinger & Arseneau, 2007). Sexual and gender minorities, then, were conceptualized as related groups based on gender non-conformity/atypicality (Alexander & Yescavage, 2003; Clarke, Hayfield, & Huxley, 2012; Drescher, 2010), as well as stigma/stress (e.g., Breslow et al., 2015; Meyer, 2015). This acknowledgment of gender-related stigma/stress is reflected in the debate regarding whether "clinically significant distress or impairment in social, occupational, or other important areas of functioning" (Criterion D in DSM-IV for Gender Identity Disorder; APA, 2000) is inherent to gender dysphoria or whether it is instead a result of social stigma, or both (Bouman, Bauer, Richards, & Coleman, 2010; Galupo, Pulice-Farrow, & Lindley, in press). It is important to note, however, that gender dysphoria does not universally apply to all individuals who identify as transgender (Byne et al., 2018), and some trans individuals do not experience dysphoria at all (Chen et al., 2016).

Given the focus on identity, research from a minority stress framework has also allowed a more nuanced understanding of transgender identity. Moving away from a focus on assigned sex and binary conceptualizations of transgender experience, recent research has broadened its understanding to be inclusive of non-binary transgender identities (Budge, Rossman, & Howard, 2014; Factor & Rothblum, 2008; Farmer & Byrd, 2015; Harrison, Grant, & Herman, 2012; Kuper, Nussbaum, & Mustanski, 2012; Lykens, LeBlanc, & Bockting, 2018; Pulice-Farrow, Clements, & Galupo, 2017; Saltzburg & Davis, 2010). Gender identity is often understood to refer to an individual's sense of self as male, female, both, or neither (Tate, 2014; Tate, Youssef, & Bettergarcia, 2014). However, recent research suggests that non-binary transgender individuals describe their gender in ways that both align with and challenge this definition (Galupo, Pulice-Farrow, & Ramirez, 2017). Acknowledgment of non-binary identities has been implicitly incorporated in the language of the most recent version of the DSM (DSM-5, Gender Dysphoria; American Psychiatric Association, 2013). Individuals are described not as identifying as "the other gender" (as was the case in the DSM-IV, Gender Identity Disorder; American Psychiatric Association, 2000) but rather as "the other gender (or some alternative gender different from one's assigned gender)" (American Psychiatric Association, 2013; Byne et al., 2018). Similar removal of the terms "opposite sex" and "anatomical sex" in earlier versions of the ICD were replaced with "experienced gender" and "assigned sex" in the ICD-11 (Reed et al., 2016; WHO, 2011).

Clinical Measures of Gender Dysphoria

Two clinical measures of gender dysphoria have been highlighted in the literature (Zucker, Lawrence, & Kreukels, 2016), both as a way to demonstrate that gender dysphoria can be reliably and validly measured (Cohen-Kettenis & Pfafflin, 2010; Singh et al., 2010) and as a way to advocate for a dimensional conceptualization of gender dysphoria (Cohen-Kettenis & Pfafflin, 2010; Deogracias et al., 2007). Following Fisk (1973, 1974), the UGDS (Cohen-Kettenis & van Goozen, 1997) conceptualized gender dysphoria on a continuum between dysphoric/not dysphoric. The scale has two non-parallel versions (female-to-male and male-to-female) and scale items focus on dissatisfaction with the body (e.g., "I hate having breasts"), gender identity (e.g., "I hate myself because I'm a boy"), and gender roles (e.g., "I prefer to behave like a boy"). In contrast, the GIDYQ-AA (Deogracias et al., 2007) was designed to provide a dimensional measure of gender identity/gender dysphoria that could be used clinically with adults and adolescents and that could be used for both trans women and men. Thus, it consists of a single questionnaire with parallel items by

---

[2] Much of the early clinical literature refers to transgender individuals on the basis of their "biological" or "natal" sex. Consistent with more contemporary understandings of transgender experience, we refer to transgender individuals by centering on their gender identity. Thus, while early typologies focused on "natal men," we describe these models as focused being focused on the experiences of trans women.

assigned sex. This scale conceptualizes gender identity/gender dysphoria on a continuum, with male and female on either pole. The GIDYQ-AA also focuses on experiences explicitly in the preceding 12-month period. Items assess subjective (e.g., "In the past 6 months, have you thought of yourself as a transgendered person?"), somatic (e.g., "In the past 12 months, have you disliked your body because it is female?"), social (e.g., "In the past 12 months, at parties or at other social gatherings, have you presented yourself as a man?"), and sociolegal (e.g., "In the past 12 months, have you made an effort to change your legal sex?") aspects of dysphoria.

Although there may be clinical utility in these two scales, the way the scales conceptualize gender dysphoria may or may not resonate with the way transgender individuals themselves experience gender dysphoria. Because the majority of research on transgender experience has been conducted from a cisgender framework, it has been suggested (Galupo, 2017) that we need to adopt research approaches that invite transgender individuals to reflect upon the theories advanced (e.g., Schudson, Dibble, & van Anders, 2017; Veale et al., 2012) and the scientific measures used (e.g., Galupo, Mitchell, & Davis, 2018; Galupo, Mitchell, Grynkiewicz, & Davis, 2014). Recently, subjective ratings have been usefully applied to sexual orientation measures as a way to explore how well they capture the experiences of both heterosexual and sexual minority individuals (Galupo et al., 2018). These findings have been used to describe whether individuals' experiences resonate with the way empirical scales are conceptualized. It is likely, then, that exploring the subjective scores for clinical scales can similarly inform current research and practice with regard to gender dysphoria.

Purpose of the Present Study

The present research utilized a non-clinical sample of transgender-identified individuals to investigate how well two clinical measures (GIDYQ-AA and UGDS) resonated with their experiences of gender dysphoria. Participants were asked to rate how well these scales captured their experience of gender dysphoria. These subjective ratings were explored across gender identity, assigned sex, and clinical diagnosis.

Clinical typologies of transsexual/transgender individuals (Benjamin, 1966; Blanchard, 1989a, 1989b) have been developed for trans women, and clinical conceptualizations of gender dysphoria have emphasized binary understandings of gender. It was, therefore, hypothesized that ratings for both scales would be higher for individuals with transfeminine versus transmasculine identities and higher for binary (transfeminine and transmasculine) versus non-binary/agender identities. It was further hypothesized that sex assigned at birth for both scales would impact subjective ratings, where individuals who were assigned male would report more favorable ratings than would individuals who were assigned female. Additionally, because these scales were both developed based on clinicians' perspectives

of working with individuals diagnosed with gender dysphoria, it was hypothesized that participants who reported having been diagnosed with either gender identity disorder or gender dysphoria would rate both scales more favorably than those who did not.

We tested our hypotheses using planned comparisons. Although we have three independent variables (gender identity, assigned sex, clinical diagnosis), planned contrasts were more appropriate to test our hypotheses than a factorial design for the following reasons: (1) we were not interested in main effects; rather, we had specific directional predictions across three discrete gender identity groups; and (2) we did not anticipate interaction effects (van Wesel & Klugkist, 2015).

Method

Participants

Participants were 622 individuals who self-identified as transgender and completed an online survey, each with a unique IP address. This analysis excluded data from participants who had incomplete or missing responses for the questions under consideration, or who indicated "not applicable" for the subjective rating question. Participants represented all regions of the USA (residing in 39 states), as well as 17 countries. Participant ages ranged from 18 to 74 years ($M = 26.89$, $SD = 8.03$). There was limited racial/ethnic diversity within the sample, with 17.4% of participants identifying as racial minorities. The overwhelming majority of the sample (82.6%) identified as White. See Table 1 for detailed information for more detailed participant demographics.

With regard to gender, participants self-selected into one of four categories: transfeminine ($n = 221$, 35.5%), transmasculine ($n = 206$, 33.1%), non-binary ($n = 150$, 24.1%), and agender ($n = 45$, 7.2%). For the purposes of our analyses, non-binary and agender groups were combined into a single non-binary/agender group. Table 2 provides demographic information across these three groups. The slight majority of the sample was assigned female at birth (58.0%). With regard to sexual identity, the vast majority of individuals (94.4%) self-identified as something other than heterosexual (asexual, bisexual, fluid, gay, lesbian, pansexual, queer, or other). Past research in a similar community sample has established that trans individuals use gender identity (not assigned sex) to guide their choice of sexual identity labels. Finally, 311 of the 622 participants reported a clinical diagnosis for either gender dysphoria or gender identity disorder.

Participants were recruited to participate in a "study on the experiences of gender dysphoria in transgender individuals." Initial recruitment announcements were posted on social networking websites and online message boards. Online resources included those with a national and local community reach. In addition, some targeted specific agender, non-binary, or

Table 1 Participant demographics

	Total ($N=622$)
Age mean (SD) (in years)	26.89 (8.79)
Race	n (%)
American Indian/Alaska Native	5 (0.8)
Asian/Asian American	13 (2.1)
Black/African-American	7 (1.1)
Hispanic/Latinx	24 (3.9)
White/Caucasian	514 (82.6)
Biracial/Multiracial	31 (5.0)
No answer	5 (0.8)
Other	23 (3.7)
Socioeconomic status	n (%)
Working class	117 (18.8)
Lower–middle class	118 (19.0)
Middle class	188 (30.2)
Upper–middle class	110 (17.7)
Upper class	6 (1.0)
Don't know	43 (6.9)
No answer	14 (2.3)
Other	26 (4.2)
Highest level of education	n (%)
High school/GED	264 (42.4)
College degree	256 (41.2)
Graduate degree	78 (12.5)
Doctorate/terminal degree	24 (3.9)

Demographics are reported in aggregate as there were no significant differences across gender identity

transgender communities of color, while others engaged the transgender community more broadly. Snowball sampling was utilized as a recruitment tool, with some participants sharing the survey with additional groups or individuals. Participants learned of the survey primarily via online resources, including seeing the survey link on Reddit (58.2%), Facebook (27.0%), Tumblr (3.9%), and Twitter (2.6%). The remaining participants were directly recruited by a friend (3.1%) or received information about the survey through another venue (5.3%). Data were collected for two months, beginning on March 26, 2018, and ending on May 26, 2018.

Measures

Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults (GIDYQ-AA)

The GIDYQ-AA (Deogracias et al., 2007) is conceptualized on a continuum between "unproblematic" gender identity to gender dysphoria. The GIDYQ-AA includes 27 items on a single-factor. Items were designed to capture subjective, social, somatic, and sociolegal aspects of gender identity/dysphoria and were expressed in parallel for women and men (based on assigned sex). To complete the scale, participants were instructed to rate the frequency of each item's occurrence within the past 12 months on a 5-point scale with 1 (*Never*) and 5 (*Always*) as anchors. The original development article reported a Cronbach's alpha of .97. For the purposes of this study, participants were presented the scale, but did not complete it. Instead, they rated the scale based on how well the scale conceptually captured their experience (see below).

Table 2 Participant demographics by gender identity

	Transfeminine $n=221$	Transmasculine $n=206$	Non-binary/agender $n=195$
Age mean (SD) (in years)	28.77 (10.35)	24.42 (7.35)	27.06 (7.81)
Assigned sex	n (%)	n (%)	n (%)
Female	0 (0.0)	206 (100.0)	154 (79.5)
Male	221 (100.0)	0 (0.0)	40 (20.5)
Clinical diagnosis	n (%)	n (%)	n (%)
Yes	123 (55.7)	129 (62.6)	59 (30.3)
No	98 (44.3)	77 (37.4)	136 (69.7)
Primary sexual identity	n (%)	n (%)	n (%)
Asexual	11 (5.0)	29 (14.1)	31 (15.9)
Bisexual	55 (24.9)	35 (17.0)	28 (14.4)
Fluid	1 (0.5)	3 (1.5)	3 (1.5)
Gay	6 (2.7)	36 (17.5)	11 (5.6)
Heterosexual	12 (5.4)	21 (10.2)	2 (1.0)
Lesbian	61 (27.6)	4 (1.9)	9 (4.6)
Pansexual	49 (22.2)	29 (14.1)	47 (24.1)
Queer	16 (7.2)	40 (19.4)	49 (25.1)
Other	10 (4.5)	9 (4.4)	15 (7.7)

Utrecht Gender Dysphoria Scale (UGDS)

The UGDS (Cohen-Kettenis & van Goozen, 1997) conceptualizes dysphoria on a continuum between not dysphoric/dysphoric. The UGDS includes 12-items where individuals rate their agreement on a 5-point scale. There are two versions of the UGDS scale that were developed independently. Cronbach's alpha has been reported as .92 for the "Male" [MTF] version and .78 for the "Female" [FTM] version.[3] For the purposes of this study, participants were presented the scale intended for their assigned sex, but did not complete it. Instead, they rated the scale based on its face validity (how well the scale conceptually captured their experience; see below).

Subjective Ratings

After reading each of the two gender dysphoria scales, participants provided their global subjective ratings for the scales. Participants responded to the following prompt: "This scale accurately reflects my experience of dysphoria." Ratings were provided using a 5-point Likert scale ranging from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). The response mode included the option to indicate N/A. Participants who indicated N/A were not included in the analysis. Total subjective ratings scores were normally distributed for both the GIDYQ-AA and UGDS and were within the normal range for all three gender identity groups.

Procedure

As part of a larger online study on gender dysphoria, participants provided basic demographic information including a self-report question regarding whether they had a previous clinical diagnoses of gender dysphoria (or gender identity disorder as it was previously known). Participants were then asked a series of open-ended questions, allowing them to describe their experiences of gender dysphoria. Finally, participants were presented with the two gender dysphoria scales without a title or label. Participants were instructed to read through the scales and then answer the subjective rating prompt.

Results

Preliminary Analysis

A preliminary analysis explored whether subjective ratings differed across individuals with non-binary ($n = 150$) and agender ($n = 45$) identities. Non-binary participants ($M = 3.19$, SD = 1.19) did not significantly differ in their subjective ratings

for the GIDYQ-AA when compared to agender participants ($M = 2.91$, SD = 1.12), $t(193) = 1.38$, *ns*. Likewise, non-binary participants ($M = 2.71$, SD = 1.15) did not significantly differ in their subjective ratings for the UGDS when compared to agender participants ($M = 2.96$, SD = 1.20), $t(193) = 1.23$, *ns*. For all analyses, then, we collapsed the two into a single non-binary/agender group ($n = 195$).

Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults

Mean subjective ratings for the GIDYQ-AA were 3.40 (SD = 1.09) on a 5-point scale. Only 52.5% of participants either agreed (38.7%) or strongly agreed (13.8%) that their experience was captured on the scale. Transfeminine ($M = 3.60$, SD = 1.02) and transmasculine ($M = 3.44$, SD = 1.03) individuals did not significantly differ in their subjective ratings for GIDYQ-AA, $t(425) = 1.57$, $p = .12$. Non-binary/agender ($M = 3.12$, SD = 1.18) individuals had significantly lower subjective scores for the GIDYQ-AA when compared to either transfeminine, $t(385.92) = 4.35$, $p < .001$, Cohen's $d = .44$ or transmasculine individuals $t(384.70) = 2.87$, $p < .01$, Cohen's $d = .29$. There was also a significant effect of sex assigned at birth where those who were assigned male ($M = 3.50$, SD = 1.05) rated the GIDYQ-AA as a more valid measure than those who were assigned female ($M = 3.32$, SD = 1.11), $t(620) = -1.97$, $p < .05$, Cohen's $d = .17$. No significant difference was found in subjective ratings based on clinical diagnosis, $t(620) = .70$, $p < .48$. See Tables 3 and 4 for means and distributions across groups.

Utrecht Gender Dysphoria Scale

Mean subjective ratings for the UGDS were 3.43 (SD = 1.22) on a 5-point scale. Only 54.6% of participants either agreed (33.1%) or strongly agreed (21.5%) that their experience was captured on the scale. Subjective ratings for the UGDS were higher for transfeminine participants ($M = 3.98$, SD = 1.11) than transmasculine ($M = 3.47$, SD = 1.10) individuals, $t(425) = 4.74$, $p < .001$, Cohen's $d = .46$ and non-binary/agender individuals ($M = 2.77$, SD = 1.16,), $t(401.51) = 10.82$, $p < .001$, Cohen's $d = 1.07$. Transmasculine individuals rated the UGDS as a more valid measure of dysphoria than did non-binary/agender individuals, $t(399) = 6.20$, $p < .001$, Cohen's $d = .62$. There was also a significant effect of sex assigned at birth where those who were assigned male ($M = 3.84$, SD = 1.17) rated the UGDS as a more valid measure than did those who were assigned female ($M = 3.14$, SD = 1.18), $t(620) = -7.29$, $p < .001$, Cohen's $d = .60$. No significant difference was found in subjective ratings based on clinical diagnosis, $t(620) = 1.38$, $p < .17$.

---

[3] Please note that in the original, the scale versions were designated based on individuals' "biological" sex and not gender identity.

Table 3 Mean subjective ratings by gender identity, assigned sex, and diagnosis

	GIDYQ-AA M (SD)	UGDS M (SD)
Gender identity		
Transfeminine	3.60 (1.02)	3.98 (1.11)
Transmasculine	3.44 (1.03)	3.47 (1.10)
Non-binary/agender	3.12 (1.18)	2.77 (1.16)
Assigned sex		
Female	3.22 (1.11)	3.14 (1.18)
Male	3.50 (1.05)	3.84 (1.17)
Clinical diagnosis		
Yes	3.37 (1.11)	3.50 (1.23)
No	3.43 (1.07)	3.36 (1.22)
Overall	3.40 (1.09)	3.43 (1.22)

Absolute range, 1–5

Subjective Ratings Across Scale

Subjective ratings for the GIDYQ-AA and UGDS scales were positively correlated for transfeminine ($r = .21$, $p < .001$), transmasculine ($r = .34$, $p < .001$), and non-binary/agender individuals ($r = .35$, $p < .001$). Although we had no a priori hypotheses regarding how subjective ratings may differ across the two scales, we conducted three pairwise comparisons to explore differences (GIDYQ-AA vs UGDS) for each of the three gender identity groups. For transfeminine participants, subjective ratings scores were significantly lower for the GIDYQ-AA ($M = 3.59$, SD $= 1.02$) when compared to the UGDS ($M = 3.98$, SD $= 1.10$), $t(220) = -4.23$, $p < .001$, Cohen's $d = .36$. For transmasculine participants, subjective ratings scores did not significantly differ between the GIDYQ-AA ($M = 3.44$, SD $= 1.03$) when compared to the UGDS ($M = 3.47$, SD $= 1.10$), $t(205) = -.341$, *ns,* Cohen's $d = .03$. For non-binary/agender, subjective ratings scores were significantly higher for the GIDYQ-AA ($M = 3.12$, SD $= 1.18$) when compared to the UGDS ($M = 2.77$, SD $= 1.16$), $t(194) = 3.72$, $p < .001$, Cohen's $d = .30$.

Direct comparisons also revealed that among participants assigned female at birth, subjective ratings were significantly higher for the GIDYQ-AA ($M = 3.32$, SD $= 1.11$) when compared to the UGDS ($M = 3.13$, SD $= 1.18$), $t(360) = 2.80$, $p < .01$, Cohen's $d = .16$. For participants assigned male at birth, subjective ratings were significantly lower for the GIDYQ-AA ($M = 3.50$, SD $= 1.05$) when compared to the UGDS ($M = 3.83$, SD $= 1.17$), $t(260) = -3.98$, $p < .001$, Cohen's $d = .30$.

Discussion

The present findings provide a way to think about how well existing scales capture transgender individuals' experience of dysphoria. Because both scales under study were developed based upon clinicians' experience and conceptualization of gender dysphoria, these findings allow us to assess how well that conceptualization resonates with transgender individuals' lived experience.

Before interpreting the group level data, a general understanding of participants' subjective ratings is gleaned when considering the overall descriptive data. It is important to note that while mean subjective scores for the two scales were relatively low (3.40 and 3.43 for the GIDYQ-AA and UGDS, respectively) more than half of the participants overall (52.5% GIDYQ-AA; 54% UGDS) agreed or strongly agreed that the scales captured their experience (versus 26.7% GIDYQ-AA; 33.1% UGDS who disagreed or strongly disagreed). Having close to half of the participants feeling neutral or negative toward the scales, however, is not without concern. The present findings call into question how well the GIDYQ-AA (Deogracias et al., 2007) and UGDS (Cohen-Kettenis & van Goozen, 1997) align with broad range of transgender individuals' experience of gender dysphoria. The comparative data allow an understanding of the pattern of responses, where transgender individuals of different identities did not equally perceive their experiences to be captured by these clinical scales of gender dysphoria.

Gender Dysphoria Measurement and Clinical Diagnosis

It has been suggested that dimensional measures of gender dysphoria could be particularly helpful for detecting non-clinical

Table 4 Subjective ratings responses across gender identity

	Transfeminine		Transmasculine		Non-binary/agender	
	GIDYQ-AA (%)	UGDS (%)	GIDYQ-AA (%)	UGDS (%)	GIDYQ-AA (%)	UGDS (%)
Strongly agree	19.5	39.4	12.6	19.4	8.7	17.4
Agree	37.1	35.7	41.7	33.0	37.4	25.6
Neutral	31.2	11.8	27.7	26.7	24.1	23.1
Disagree	8.1	9.5	13.1	17.0	16.9	30.3
Strongly disagree	4.1	3.6	4.9	3.9	12.8	3.6

levels dysphoria (Cohen-Kettenis & Pfafflin, 2010; Deogracias et al., 2007). The majority of the published research using these scales, however, have utilized clinical samples (i.e., where transgender samples consisted of clients recruited from gender clinics). The present research used an online sample of transgender individuals, yielding a diverse sample with regard to clinical experiences. Exactly half of the sample reported having a clinical diagnosis of either gender identity disorder or gender dysphoria. Contrary to our hypothesis, subjective ratings did not differ across clinical diagnosis. These findings suggest that participants with a clinical diagnosis were not more likely to feel the scales captured their experience of dysphoria than those participants without a diagnosis of dysphoria. The relatively low subjective ratings of these scales, then, cannot be attributed to the heterogeneity of the sample with regard to clinical experiences. Caution should be used when interpreting these scores in a clinical setting as a definitive measure of dysphoria. The scale responses, instead, could be used as a starting point for clinicians in understanding the way gender dysphoria is being experienced by an individual. For example, a client's answers on these measures may serve as a helpful prompt for a conversation regarding which items are salient to their dysphoria, and what facets of their dysphoria are not included that they may want to address in therapy.

Gender Dysphoria Measurement across Gender Identity and Assigned Sex

We hypothesized differences in subjective ratings across gender identity and assigned sex for both measures of gender dysphoria. Our findings supported these hypotheses. Participants who were assigned male rated the scales to be a more valid measure of their experiences than those assigned female. This is not altogether surprising as trans women (i.e., transfeminine individuals who were assigned male) have long been the focus in the literature. However, this finding has important implications for how scores on these measures should be interpreted. When using both scales, researchers should contemplate whether the inherent bias in the measurement used might impact the conclusions they are drawing from their findings. For example, many researchers have reported that when assessed using these two scales, trans men (i.e., transmasculine individuals who were assigned female) report higher levels of dysphoria than trans women (Cohen-Kettenis & van Goozen, 1997; Deogracias et al., 2007; Schneider et al., 2016; Singh et al., 2010). However, our findings suggest that caution should be used when interpreting these putative gender differences, as individuals who were assigned female also rate these scales to be a less accurate measure of their dysphoria.

Differences in subjective ratings were also found across gender identity. For the GIDYQ-AA, transfeminine and transmasculine individuals rated the scales to be a more accurate reflection of their dysphoria than did non-binary/agender individuals.

For the UGDS, transfeminine individuals reported the highest subjective scores, followed by transmasculine individuals, and then non-binary/agender individuals. These findings followed our predicted patterns and suggest that for both scales, non-binary and agender individuals report the lowest subjective scores. These scales, then, are more conceptually aligned with binary experiences of gender. The UGDS also demonstrated lower subjective ratings for transmasculine individuals when compared to transfeminine individuals, suggesting that the UGDS is most aligned with the experiences of trans women in particular. This is supported by direct comparisons where transfeminine individuals rated the UGDS more favorably than the GIDYQ-AA, while non-binary/agender participants favored the UGDS. No differences in subjective ratings between the two scales were found for transmasculine individuals.

Both dysphoria scales were framed to consider gender dysphoria on the basis of an individual's assigned sex (Cohen-Kettenis & van Goozen, 1997; Deogracias et al., 2007). For example, both scales have two versions—one for men and another for women. For the GIDYQ-AA, each item is worded in parallel across the two versions, allowing for direct comparisons. For the UGDS, the two versions are distinct. Although subjective ratings differed across assigned sex and gender identity for both scales, effect sizes were larger for the UGDS. Direct comparisons demonstrated participants assigned female favored the GIDYQ-AA, while those assigned male rated the UGDS more favorably. The present findings support the need to investigate the ways that gender dysphoria may be experienced differently across gender identity and to develop measures that reflect unique experiences that are informed by both gender identity and assigned sex.

Limitations and Directions for Future Research

Our participants represented a convenience sample collected online. Online recruitment is particularly useful for transgender research where participants may have privacy concerns and may not otherwise have access for participation (Riggle, Rostosky, & Reedy, 2005). It was also useful as we were targeting a non-clinical sample. Online sampling, however, disproportionately represents educated, middle class, White individuals (Dillman, Smyth, & Christian, 2008). With 82.6% of our participants identifying as White and 48.9% identifying as middle or upper class, our sample demographics were consistent with this trend. Thus, interpretation of these data should be noted within the sample demographics. Our findings further indicated that non-binary/agender individuals rated both scales as less likely to capture their experience of dysphoria than the other groups. It is important to note that although our participants all answered a recruitment for call for a "study on the experiences of gender dysphoria in transgender individuals," this language may not have resonated with all non-binary/agender individuals.

Our findings suggest that subjective ratings did not significantly differ based on clinical diagnosis. That is, those with a
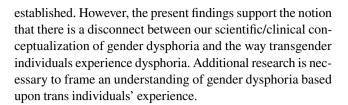
clinical diagnosis of gender dysphoria were not more likely to rate the two scales as salient to their gender dysphoria than those without a diagnosis. It is important to note, however, that clinical diagnosis was determined via self-report and may not be as accurate as data derived from medical records. Another limitation is that we do not have transition status data available due to an error in the way those variables were formatted in the survey platform for data collection. It will be important for future research to consider how transition status impacts the subjective ratings of clinical scales for gender dysphoria.

The present research utilized subjective ratings as a way to assess how well two clinical scales resonate with trans individuals' lived experience of gender dysphoria. Overall, these findings suggest that trans individuals do not feel either scale captures their experiences very well. Although they were not rated to be a very strong representation of their experience of dysphoria, participants' subjective ratings were based on their global assessment of the scale and not on individual items. Subjective ratings, then, did not provide nuanced information about scale items and did not provide insight into which specific aspects of their dysphoria were captured by these scales, and which aspects were left out. Future qualitative research will be necessary to better understand this. Such qualitative research should focus on identifying factors that contribute to each scales' subjective ratings and should explore whether salient factors differ across gender identity and assigned sex. It is likely that this information could be used to develop scales that better resonate with transgender individuals' experiences of gender dysphoria.

Implications for Measurement of Gender Dysphoria

When using traditional measures of gender dysphoria, researchers should contemplate whether the inherent bias in the measurement might impact the conclusions they are drawing from their findings. Likewise, researchers should consider whether their current approach to collecting demographic information from their sample is sufficient. For example, many clinical studies of dysphoria primarily consider assigned sex (over gender identity). When information about gender identity is collected, questions should include non-binary trans identities. Because subjective ratings were lowest for non-binary/agender individuals, it will be important for researchers and clinicians to collect gender identity information and consider this information when interpreting gender dysphoria scores.

The present study answers the call for more research approaches that invite transgender individuals to reflect upon the theories advanced and measures used to better explain their experience (Galupo, 2017). Conceptualized from the experience of clinicians, the psychometrics of both the GIDYQ-AA (Deogracias et al., 2007; Singh et al., 2010) and the UGDS (Cohen-Kettenis & van Goozen, 1997) have been well

established. However, the present findings support the notion that there is a disconnect between our scientific/clinical conceptualization of gender dysphoria and the way transgender individuals experience dysphoria. Additional research is necessary to frame an understanding of gender dysphoria based upon trans individuals' experience.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Alexander, J., & Yescavage, K. (2003). Bisexuality and transgenderism. *Journal of Bisexuality, 3*(3–4), 1–23. https://doi.org/10.1300/j159v03n03_01.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Press.

Benjamin, H. (1966). *The transsexual phenomenon*. New York: Julian Press.

Blanchard, R. (1989a). The concept of autogynephilia and the typology of male gender dysphoria. *Journal of Nervous and Mental Disease, 177,* 616–623. https://doi.org/10.1097/00005053-198910000-00004.

Blanchard, R. (1989b). The classification and labeling of nonhomosexual gender dysphorias. *Archives of Sexual Behavior, 18,* 315–324. https://doi.org/10.1007/BF01541951.

Bockting, W. O., Coleman, E., & Lief, H. (1991). A comment on the concept of transhomosexuality, or the dissociation of the meaning. *Archives of Sexual Behavior, 20*(4), 419–421. https://doi.org/10.1007/BF01541951.

Bouman, W. P., Bauer, G. R., Richards, C., & Coleman, E. (2010). World Professional Association for Transgender Health Consensus Statement on Considerations of the Role of Distress (Criterion D) in the DSM diagnosis of Gender Identity Disorder. *International Journal of Transgenderism, 12*(2), 100–106. https://doi.org/10.1080/15532739.2010.513927.

Breslow, A. S., Brewster, M. E., Velez, B. L., Wong, S., Geiger, E., & Soderstrom, B. (2015). Resilience and collective action: Exploring buffers against minority stress for transgender individuals. *Psychology of Sexual Orientation and Gender Diversity, 2*(3), 253–265. https://doi.org/10.1037/sgd0000117.

Budge, S. L., Rossman, H. K., & Howard, K. A. S. (2014). Coping and psychological distress among genderqueer individuals: The moderating effect of social support. *Journal of LGBT Issues in Counseling, 8*(1), 95–117. https://doi.org/10.1080/15538605.2014.853641.

Byne, W., Karasic, D. H., Coleman, E., Eyler, A. E., Kidd, J. D., Meyer-Bahlburg, H. F. L., & Pula, J. (2018). Gender dysphoria in adults: An overview and primer for psychiatrists. *Transgender Health*. https://doi.org/10.1089/trgh.2017.0053.

Chen, D., Hidalgo, M. A., Leibowitz, S., Leininger, J., Simons, L., Finlayson, C., & Garofalo, R. (2016). Multidisciplinary care for gender-diverse youth: A narrative review and unique model of gender-affirming care. *Transgender Health, 1*(1), 117–123.

Clarke, V., Hayfield, N., & Huxley, C. (2012). Lesbian, gay, bisexual and trans appearance and embodiment: A critical review of the psychological literature. *Psychology of Sexualities Review, 3*(1), 51–70.

Cohen-Kettenis, P. T., & Pfafflin, F. (2010). The DSM diagnostic criteria for gender identity disorder in adolescents and adults. *Archives of Sexual Behavior, 39,* 499–513. https://doi.org/10.1007/s10508-009-9562-y.

Cohen-Kettenis, P. T., & van Goozen, S. H. M. (1997). Sex reassignment of adolescent transsexuals: A follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 36,* 263–271. https://doi.org/10.1097/00004583-199702000-00017.

Deogracias, J. J., Johnson, L. L., Meyer-Bahlburg, H. F. L., Kessler, S. J., Schober, J. M., & Zucker, K. J. (2007). The Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults. *Journal of Sex Research, 44,* 370–379. https://doi.org/10.1080/00224490701586730.

Devor, H. (1993). Sexual orientation identities, attractions, and practices of female-to-male transsexuals. *Journal of Sex Research, 30*(4), 303–315. https://doi.org/10.1080/00224499309551717.

Dillman, D., Smyth, J., & Christian, L. M. (2008). *Internet, mail, and mixed-mode surveys: The tailored design method*. New York: Wiley.

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. https://doi.org/10.1007/s10508-009-9531-5.

Drescher, J. (2015). Queer diagnoses revisted: The past and future of homosexuality and gender diagnoses in DSM and ICD. *International Review of Psychiatry, 27*(5), 386–395. https://doi.org/10.3109/09540261.2015.1053847.

Drescher, J., Cohen-Kettenis, P., & Winter, S. (2012). Minding the body: Situating gender identity diagnoses in the ICD-11. *International Review of Psychiatry, 24*(6), 568–577. https://doi.org/10.3109/09540261.2012.741575.

Factor, R. J., & Rothblum, E. D. (2008). Exploring gender identity and community among three groups of transgender individuals in the United States: MTFs, FTMs, and genderqueers. *Health Sociology Review, 17,* 241–241. https://doi.org/10.5172/hesr.451.17.3.235.

Farmer, L. B., & Byrd, R. (2015). Genderism in the LGBTQQIA Community: An interpretative phenomenological analysis. *Journal of LGBT Issues in Counseling, 9*(4), 288–310. https://doi.org/10.1080/15538605.2015.1103679.

Fassinger, R. E., & Arseneau, J. R. (2007). 'I'd rather get wet than be under the umbrella': Differentiating the experiences and identities of lesbian, gay, and transgender people. In K. J. Biesche, R. M. Perez, & K. A. DeBord (Eds.), *Handbook of counseling and psychotherapy with lesbian, gay, bisexual, and transgender clients* (pp. 19–49). Washington, DC: American Psychological Association.

Fisk, N. M. (1973). Gender dysphoria syndrome (The how, what and why of a disease). In D. R. Laub & P. Gandy (Eds.), *Proceedings of the second interdisciplinary symposium on gender dysphoria syndrome* (pp. 7–14). Palo Alto, CA: Stanford University Press.

Fisk, N. M. (1974). Gender dysphoria syndrome: The conceptualization that liberalizes indications for total gender reorientation and implies a broadly based multi-dimensional rehabilitative regimen. *Western Journal of Medicine, 120,* 386–391.

Galupo, M. P. (2017). Researching while cisgender: Identity considerations for transgender research. *International Journal of Transgenderism, 18*(3), 241–242. https://doi.org/10.1080/15532739.2017.1342503.

Galupo, M. P., Mitchell, R. C., & Davis, K. S. (2018). Face validity ratings of sexual orientation scales by sexual minority adults: Effects of sexual orientation and gender identity. *Archives of Sexual Behavior, 47,* 1241–1250. https://doi.org/10.1007/s10508-017-1037-y.

Galupo, M. P., Mitchell, R. C., Grynkiewicz, A. L., & Davis, K. S. (2014). Sexual minority reflections on the Kinsey Scale and the Klein Sexual Orientation Grid: Conceptualization and measurement. *Journal of Bisexuality, 14,* 414–432. https://doi.org/10.1080/15299716.2014.929553.

Galupo, M. P., Pulice-Farrow, L., & Lindley, L. (in press). "Every time I get gendered male, I feel a pain in my chest": Understanding the social context for gender dysphoria. *Stigma & Health.* https://doi.org/10.1037/sah0000189.

Galupo, M. P., Pulice-Farrow, L., & Ramirez, J. L. (2017). "Like a constantly flowing river": Gender identity flexibility among non-binary transgender individuals. In J. D. Sinnott (Ed.), *Identity flexibility during adulthood: Perspectives in adult development* (pp. 163–177). Cham, Switzerland: Springer.

Harrison, J., Grant, J., & Herman, J. L. (2012). A gender not listed here: Genderqueers, gender revels, and otherwise in the National Transgender Discrimination Survey. *LGBTQ Policy Journal at the Harvard Kennedy School, 2,* 13–24. Retrieved from: http://escholarship.org/uc/item/2zj46213.

Hendricks, M. L., & Testa, R. J. (2012). A conceptual framework for clinical work with transgender and gender nonconforming clients: An adaptation of the minority stress model. *Professional Psychology: Research and Practice, 43*(5), 46–467. https://doi.org/10.1037/a0029597.

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,* 244–254. https://doi.org/10.1080/00224499.2011.596954.

Lykens, J. E., LeBlanc, A. J., & Bockting, W. O. (2018). Healthcare experiences among young adults who identify as genderqueer or nonbinary. *LGBT Health, 5,* 191–196. https://doi.org/10.1089/lgbt.2017.0215.

Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin, 129,* 674–697.

Meyer, I. H. (2015). Resilience in the study of minority stress and health of sexual and gender minorities. *Psychology of Sexual Orientation and Gender Diversity, 2*(3), 209–213. https://doi.org/10.1037/sgd0000132.

Meyerowitz, J. (2002). *How sex changed: A history of transsexuality in the United States*. Cambridge, MA: Harvard University Press.

Pflum, S. R., Testa, R. J., Balsam, K. F., Goldblum, P. B., & Bongar, B. (2015). Social support, trans community connectedness, and mental health symptoms among transgender and gender nonconforming adults. *Psychology of Sexual Orientation and Gender Diversity, 2*(3), 281–281. https://doi.org/10.1037/sgd0000122.

Pulice-Farrow, L., Clements, Z. A., & Galupo, M. P. (2017). Patterns of transgender microaggressions in friendship: the role of gender identity. *Psychology & Sexuality, 8*(3), 189–207. https://doi.org/10.1080/19419899.2017.1343745.

Reed, G. M., Drescher, J., Krueger, R. B., Atalla, E., Cochran, S. D., First, M. B., & Saxena, S. (2016). Disorders related to sexuality and gender identity in the ICD-11: Revising the ICD-10 classification based on current scientific evidence, best clinical practices, and human rights considerations. *World Psychiatry, 15*(3), 205–221. https://doi.org/10.1002/wps.20354.

Riggle, E. D. B., Rostosky, S. S., & Reedy, C. S. (2005). Online surveys for BGLT research. *Journal of Homosexuality, 49,* 1–21. https://doi.org/10.1300/j082v49n01-01.

Saltzburg, S., & Davis, T. S. (2010). Co-authoring gender-queer youth identities: Discursive tellings and retellings. *Journal of Ethnic and Cultural Diversity in Social Work, 19*(2), 87–108. https://doi.org/10.1080/15313200903124028.

Schneider, C., Cerwenka, S., Nieder, T. O., Briken, P., Cohen-Kettenis, P. T., De Cuypere, G., … Richter-Appelt, H. (2016). Measuring gender dysphoria: A multicenter examination and comparison of the Utrecht Gender Dysphoria Scale and the Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults. *Archives of Sexual Behavior, 45*(3), 551–558. https://doi.org/10.1007/s10508-016-0702-x.

Schudson, Z., Dibble, E. R., & van Anders, S. M. (2017). Gender/sex and sexual diversity via sexual configurations theory: Insights from a qualitative study with gender and sexual minorities. *Psychology of Sexual Orientation and Gender Diversity, 4*(4), 422–437.

Serano, J. M. (2010). The case against autogynephilia. *International Journal of Transgenderism, 12*(3), 176–187. https://doi.org/10.1080/15532739.2010.514223.

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. https://doi.org/10.1080/00224490902898728.

Tate, C. C. (2014). Gender identity as a personality process. In B. L. Miller (Ed.), *Gender identity: Disorders, developmental perspectives, and social implications* (pp. 1–22). Hauppauge, NY: Nova Science Publishers.

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. https://doi.org/10.1037/gpr0000019.

Timmins, L., Rimes, K. A., & Rahman, Q. (2017). Minority stressors and psychological distress in transgender individuals. *Psychology of Sexual Orientation and Gender Diversity, 4*(3), 328–340. https://doi.org/10.1037/sgd0000237.

van Anders, S. M. (2015). Beyond sexual orientation: Integrating gender/sex and diverse sexualities via sexual configurations theory. *Archives of Sexual Behavior, 44,* 1177–1213. https://doi.org/10.1007/s10508-015-0490-8.

van Wesel, F., Klugkist, I., & Hoijtink. H. (2015). *Formulating and evaluating interaction effects*. https://informative-hypotheses.sites.uu.nl/wp-content/uploads/sites/23/2015/04/van-wesel-klugkist-and-hoijtink.pdf.

Veale, J. F., Clarke, D. E., & Lomax, T. C. (2012). Male-to-female transsexuals' impressions of Blanchard's autogynephilia theory. *International Journal of Transgenderism, 13*(3), 131–139. https://doi.org/10.1080/15532739.2011.669659.

World Health Organization. (2011). *The ICD-11 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Zucker, K. J., Lawrence, A. A., & Kreukels, B. P. C. (2016). *Annual Review of Clinical Psychology, 12,* 217–247. https://doi.org/10.1146/annurev-clinpsy-021815-093034.