The Genderqueer Identity (GQI) Scale: Measurement and validation of four distinct subscales with trans and LGBQ clinical and community samples in two countries

Jenifer K. McGuire, Titia F. Beek, Jory M. Catalpa & Thomas D. Steensma

Published online: 25 May 2018.

Submit your article to this journal ⬀

View related articles ⬀

View Crossmark data ⬀

Taylor & Francis
Taylor & Francis Group

Check for updates

The Genderqueer Identity (GQI) Scale: Measurement and validation of four distinct subscales with trans and LGBQ clinical and community samples in two countries

Jenifer K. McGuire[a], Titia F. Beek[b], Jory M. Catalpa[a], and Thomas D. Steensma[b]

[a]University of Minnesota, St. Paul, Minnesota, U.S; [b]VU University Medical Center, Department of Medical Psychology, Center of Expertise on Gender Dysphoria, Amsterdam, the Netherlands

ABSTRACT

Background: Non-binary gender measurement has grown out of a need for accurate representation in scholarship and public health services available to a diverse gender population.

Aims: The Genderqueer Identity Scale (GQI) was developed to allow for a multidimensional assessment of genderqueer identity, including non-binary identity, socially constructed versus essentialist gender, theoretical awareness of gender concepts, and gender fluidity. The GQI was designed to assess gender identity across a full spectrum of gender, at any age after mid-adolescence, and at various stages of gender identity development, including prior to, during, and after a gender transition, where applicable. Two of the GQI subscales focus on intrapersonal processes, while two focus on interpersonal processes.

Methods: The measure was piloted and refined across four distinct samples: a U.S. university based LGBT sample, consecutive clinical referrals at the Center of Expertise on Gender Dysphoria in Amsterdam, the Netherlands, a Dutch LGB community sample, and an online survey forum (LGBTQ).

Results: The first exploratory factor analysis identified minor potential adjustments, which were refined and retested. Researchers evaluated and cross-validated the hypothesized factor structure and determined that the three factor GQI subscales and the unidimensional Gender Fluidity measure yielded internally consistent and valid scores among transgender individuals seeking clinical treatment and LGB individuals within a community setting. The exploratory and confirmatory factor analyses provide evidence of good reliability, construct validity, and internal consistency of all four subscales.

Discussion: The subscales were appropriate across a spectrum of gender identities and can be taken in the same form over time and across gender transition statuses, making them suitable for clinical evaluation and community based longitudinal research with trans-identified or gender nonconforming persons. The development of the GQI fills critical gaps in gender-related measurement including the ability to assess multiple dimensions of gender identity, and to assess gender identity across time.

Introduction

Gender has historically been dominated by a social binary in both theory and measurement (Bauer et al., 2009). In this article, we seek to re-center the latent concept and measurement of gender identity around a genderqueer norm, with domains representing the multiple embodied, social, political, and temporal contributors. Our goal is to understand and measure gender in a way that captures gender role as well as gender expression, supports diagnostic needs for transgender health services providing gender affirming medical interventions, including hormone treatment and surgery, as well as measurement in populations or studies of gender (Reisner et al., 2015). We begin with some fundamental assumptions about gender and measurement, move to a theoretical overview that guided the development of a new measure of genderqueer identity, and collection of data which validates the new measure across samples from several population groups. We next provide background on the relevant data history and why empirically this measure was necessary at this time. Finally, we present data from a series of pilot studies used to validate the measure on a variety of systematically chosen subgroups.

Numerous innovations in research with transgender spectrum individuals have highlighted multiple dimensions of gender identity development and the diversity of transgender and genderqueer individuals (Bockting, 2014; Diamond, Pardo, & Butterworth,

2011; Egan & Perry, 2001; Richards, Bouman, & Barker, 2017). Scholars have called for population-based studies that accurately represent and affirm transgender and genderqueer people's identities and needs (dickey, Hendricks, & Bockting, 2016; Rankin & Garvey, 2015; Reisner et al., 2015). For example, Reisner et al. (2015) recommended that researchers measure multiple aspects of gender, offering response options that go beyond binary categories. dickey et al. (2016) accurately expressed need for a specific focus on transgender individuals and challenged the promotion of census question types that erased some transgender persons' identities, such as people who do not identify as transgender, but are of transgender experience (i.e., people who have transitioned and identify as male or female and some genderqueer individuals). Similarly, Rankin and Garvey (2015) called for researchers to include multidimensional measures of gender identity that capture nuanced aspects of a person's identity. Scholars' requests for nuanced gender identity measurement underpin the construction of the Genderqueer Identity Scale (GQI).

Genderqueer or non-binary identities represent individuals who do not subscribe to the conventional gender binary and, instead, identify as *both* masculine and feminine, *between* masculine or feminine, gender neutral, or outside of the gender binary (Diamond et al., 2011; Kozee, Tylka, & Bauerband, 2012; Richards et al., 2016). Genderqueer and non-binary individuals are often subsumed under lesbian, gay, bisexual, queer, and transgender umbrellas (LGBTQ), with little to no focus on nuances specific genderqueer identity development (Fassinger & Arseneau, 2007; Harrison, Grant, & Herman, 2012). Researchers found in a sample of 6,456 transgender identified participants 13.3% said their gender was not listed among the given options of male, female, or transgender (Grant et al., 2011; Harrison et al., 2012). Authors analyzed the responses the participants wrote in finding that the majority self-defined as "genderqueer or some variation thereof, such as pangender, third gender, or hybrid" (Harrison et al., 2012).

Further, gender identity development trajectories vary widely within the transgender spectrum population (Bockting, 2014). Historically, scholarship has focused solely on transgender individuals seeking medical transition, while negating the medical needs of individuals who did not present at clinics for psychotherapy or gender affirming medical interventions, but who none-the-less identified with a gender that differed from their sex assigned at birth (Diamond et al., 2011). Gender-related measurement that extends beyond traditional, binary representations of masculinity and femininity has grown out of a need for accurate representation in scholarship and adequate public health services available to a diverse gender population. Bem (1974) validated a 60-item measure of personality traits associated with gender characteristics. One of the first measures to assess variable gender, a total score on the Bem Sex Role Inventory (BSRI) was purported to correlate with one's gender identity (i.e., masculinity, femininity, androgyny, or undifferentiated).

Conceptual framework

We draw from a broad range of theoretical frameworks and research and clinical experiences to develop, test, and interpret this measure. Each of the subscales was developed by an integrated team of people, and incorporates biopsychosocial, evolutionary, intersectional, feminist, communication, social learning, and queer theoretical perspectives. Queer theory provides a foundational perspective for this work (Nagoshi, Brzuzy, & Terrell, 2012; Serano, 2007). Fundamentally, gender identity is viewed as a latent construct rather than a single manifest variable (male, female, genderqueer). Genderqueer is at the center of the construct, and multiple factors contribute to levels of genderqueer identification. The re-centering of a long-time binary on a spectrum shifts the overall concept of gender from an either/or to relative levels of different contributors. Stryker (2008) posited that queer was used to describe different kinds of people coming together in their opposition of oppression. Thus, genderqueer can represent individuals coming together in resistance to the construct of gender as a fixed, biologically determined, natural binary of male and female. Some features of queer theory that are specifically relevant in this work include the active construction of the self, embodiment of identity, and transgression of social norms as an element of self-construction (Butler, 1990; Wilchins, 2004). For each subscale, unique theoretical perspectives contributed to its development. Below we describe the conceptual frame that orients each subscale.

Challenge the binary. One feature of genderqueer identity is the extent to which an individual challenges the gender binary, and how much that challenge is

intentional. People are socialized to enact gendered behavior in accordance with binary gender norms (Butler, 1990; Johnson, 2016; West & Zimmerman, 1987). Some people resist binary expressions and present a blended masculine and feminine identity and others actively construct genderqueerness as an intentional transgression of gendered social norms. Intentionally altering one's body or appearance in an effort to counter prevailing social expectations of a gender binary represents a component of gender identity that is very different from a person who appears non-binary. Our first domain in genderqueer identity isolates the extent to which someone is non-binary in expression, and feels that their presentation and behavior has an intentionally transgressive component.

Socially constructed gender. For decades, the nature of gender as a social construct has dominated academic and political discourse about gender roles, gender performance, gender and power and human rights (Butler, 1990; Nagoshi & Brzuzy, 2010; Wilchins, 2004). Tensions about the nature of sex, gender identity, and gender expression are often reduced to dialectics, with one camp arguing for a socially constructed view of gender and the other arguing these characteristics are largely biologically driven (Nagoshi & Brzuzy, 2010; Nagoshi et al., 2012). In the case of gender nonconformity, an excessive reliance on social construction as a developmental force for gender identity has revealed itself to be brutally inhumane (Money, 1975). Feminist, queer, and transgender theorists have come to reconsider the driving force and empirical research on trans persons has made clear that for many, although certainly not all, gender identity and expression was something that seemed to emanate from within them, and was immutable to the many (sometimes aggressive) attempts of their parents, teachers, and clergy to change (Nagoshi et al., 2012; Serano, 2007). Overtime, it has become increasingly clear in theory and research that individuals vary in the extent to which they actively construct or perform their gender versus the extent to which their gender simply is and always was a certain way (McGuire, Kuvalanka, Catalpa, & Toomey, 2016; Nagoshi & Brzuzy, 2010; Pollock & Eyre, 2012). Both discourses accurately explain aspects of gender identity development and fail to account for other aspects (McGuire et al., 2016). For example, scholars have successfully argued that gender identity is learned through socialization (West &

Zimmerman, 1987) and enacted through performance and social interaction (Butler, 1990). However, to conceptualize gender as *only* socially constructed negates aspects of gender that are felt deeply and internally and endured despite socialization efforts to eradicate such behaviors (Serano, 2007).

Theoretical awareness. Embodiment of identity is another essential aspect of queer theory that informs our work. Embodiment is the concept that our physical selves represent our political or social identity that we mark ourselves in some way (Butler, 1988). Active embodiment of gender includes the altering of one's body and dress to meet political or social goals, and the awareness of one's body as a political statement of identity; "the personal is political" (Butler, 1988; Stryker, 2008). For genderqueer persons, and persons who are intellectually sensitive to non-binary gender identities, making space for gender diversity among others is one way to make personal identity political, to embody gender. Our subscale of theoretical and political awareness is one of the scales where someone who may be *cis*appearing, or *cis*gender, may still endorse the items because they intellectually function in genderqueer theoretical and political domains such as reading, critical reflection, and political action concerning how their personal identity is situated in political systems of heterosexism, transphobia, and misogyny (Serano, 2007). This subscale is part of the latent construct that allows for a broad conceptualization of gender that is not simply limited to appearance, but also includes the ways in which someone reports interacting with the gender of others. Political issues are an aspect of gender identity where nuanced subjective experiences intersect, possibly through common experiences with *cis*normative oppression (Bauer et al., 2009).

Fluidity. Fluidity differs significantly from non-binary identity in that fluidity is about the frequency of shifting back and forth in gender expression, identity, and experience. Fluidity is not primarily concerned with the one-time process of transition in a migration-like pattern (Diamond et al., 2011), rather more of an oscillation or boundary negation where there is an intention of blurring the boundary or moving back and forth across the boundaries between gender identities (Diamond & Butterworth, 2008). Fluidity contains a transgressive element, like non-binary identity, because it involves intentional change across boundaries back and forth overtime and in different contexts (Harrison

et al., 2012). A gender fluid person may present in a fluid fashion most of the time with multiple expressions blending elements of different gender identities, or may fall closer to a bigender identity, sometimes expressing more male and other times more female identity characteristics. All iterations of fluidity are likely predominately non-binary, although iterations of non-binary identities may not be fluid (e.g., stable gender neutral presentation).

Existing binary measures

Since 1990, the Recalled Gender Identity Scale (RCGI) has been commonly used, with cross validation in a variety of populations (Zucker et al., 2006). The 23-item measure assessed childhood gender identity, gender role, and feelings about one's same and other gendered parents and siblings. (Zucker et al., 2006). The scale uses separate male and female versions that reinforce a gender binary. For instance, on both versions one question reads "as a child, I felt…" with a 5-point Likert-type scale response pattern ranging from very masculine to very feminine, or not masculine or feminine. While the question is the same on both versions, there is no way for a respondent to answer both masculine and feminine, thus reinforcing a gender binary where gender is dichotomized and distinct.

A variety of measurement problems are introduced when individuals shift versions of an instrument over time, especially if the two separate versions lack parity. First, having distinct versions implies that a distinct change from one gender to another happens in a binary fashion, allowing little or no "transition time." Historically, transition referred to gender affirming medical interventions (GCMIs; e.g., hormone treatment and gender-related surgery); however, contemporarily transition has been used to describe the process of change from one's assigned sex at birth to an identity under the transgender or genderqueer umbrellas (e.g., trans, transfemale, transmale, transfeminine, transmasculine, genderqueer, genderfluid, non-binary, or agender) (Bockting, 2014; Budge et al., 2013). A social transition occurs when an individual takes on social roles more commonly applied to a gender different than the one they were assigned at birth, regardless of whether or not medical intervention is sought (Olson, Durwood, DeMeules, & McLaughlin, 2016). In practice, the transition process can take months or years, and adjustments to gender identity

and gender roles do not simply switch from one set of expectations to another.

Second, measurement error is exacerbated when the orientation of a scale is shifted. For instance, a person who scores low on masculine identity (as an assigned male at birth) may or may not score high on femininity as an affirmed female. Affirmed female here refers to a transwoman, assigned male at birth and affirmed as a female in her current gender identity. From a measurement perspective, the mean on one scale at time one should not be compared to the mean on a different scale at time two. Finally, the error variance of items should be retained across time points, and when individuals take a different version of a scale, different items will have different error variance, thus limiting the validity of the comparison.

Similarly, two other widely used measures of gender dysphoria used two versions, one for transfemale and another for transmale individuals (Cohen-Kettenis & van Goozen, 1997; Deogracias et al., 2007). The Utrecht Gender Dysphoria Scale (UGDS) is a 12-item subscale that focuses on dissatisfaction with bodily aspects of gender identity or gender role, in the moment; whereas, the Gender Identity/Gender Dysphoria Questionnaire for Adolescents and Adults (GIDYQ-AA) measures subjective and sociological aspects of gender identity and feelings of gender dysphoria, within the past year (Deogracias et al., 2007; Schneider et al., 2016; Singh et al., 2010). Scholars administered both measures in a study of 318 transgender individuals recruited from four European gender clinics, finding that transmale participants reported stronger gender dysphoria than the transfemale participants, stronger dissatisfaction toward their bodies in the UGDS, and more distress about their gender identity in the GIDYQ-AA (Schneider et al., 2016). Scholars reported that both measures were strongly correlated for male-to-female participants, while for female-to-male participants the UGDS had ceiling effects, suggesting there are constraints with the UGDS to gather sufficient data for transpersons assigned female at birth.

The fact that the RCGI, UGDS, and GIDYQ-AA all use two separate versions of the instrument to be administered based on assigned sex at birth represents a significant limitation for each measure. In all three measures, the male version is worded slightly different from the female version limiting researchers' ability to administer the measures longitudinally. For example, the RCGI

asks assigned males about having a reputation as a "sissy" and assigned females about having a reputation as a "tomboy." Similarly, the UGDS-Female version prompts, "I wish I had been born as a boy/man," whereas, the UGDS-Male version prompts, "it would be better not to live than to live as a boy/man" (Cohen-Kettenis & van Goozen, 1997). The measures lose their applicability after transition and it becomes unclear how clinicians and researchers best implement the measure post-transition (e.g., are transwomen given the male version of the UGDS after transitioning to female). Moreover, because each instrument uses a male version and a female version, the measures implicitly reinforce the gender binary and do not adequately assess non-binary and genderqueer identities. However, the GIDYQ-AA does include one question on both the male and female versions that may elicit non-binary gender type response, asking, "In the past 12 months, have you felt uncertain about your gender, that is, feeling somewhere between a man and a woman?" Although, this single question is inadequate to assess and classify genderqueer/non-binary populations.

Proper assessment and classification of gender identity and expression has an important public health impact. Historically, transgender and gender diverse persons had to follow a master narrative following "a binary model of gender as either male or female, man or woman, masculine or feminine, with 'true' transsexuals reporting that they had been born in the wrong body" (dickey et al., 2016). These implicit biases are built into the previously discussed measures that limit researchers' capacity to acquire an inclusive count of vulnerable and erased genderqueer/non-binary populations. Accurate estimates are needed to show a public presence, thus a public need, which can be used to leverage much needed public and private resources (Doan, 2016).

As Rosario, Schrimshaw, Hunter, and Levy-Warren (2009) has argued, gender identity affects developmental trajectories and public health needs. In a study of 386 adults who applied for gender identity treatment in the Netherlands, 26.9% of the participants requested partial treatment transition-related services (e.g., requests for hormones and chest reconstructive surgery ("top surgery"), but not phalloplasty, metoidioplasty or vaginoplasty, commonly referred to as "bottom surgery") (Beek, Kreukels, Cohen-Kettenis, & Steensma, 2015). Most requested partial treatment due to concerns about risks associated with invasive and irreversible surgical procedures, or lack of genital dysphoria, while only 4% explicitly self-identified as non-binary. Participants assigned female at birth were more likely than participants assigned male at birth to request partial treatment, and endorse surgical risk; whereas, male assigned were more likely to endorse a lack of genital dysphoria as motivation for partial treatment (Beek et al., 2015).

In this article, we introduce the GQI, a questionnaire we developed to measure genderqueer and non-binary identification and expression longitudinally, across a broad spectrum, before, during, or after transition. The sub-constructs were developed by incorporating existing measurement constructs, existing theory on gender, trans identities, and genderqueer identity, qualitative interview data from the lead researcher, and extensive clinical experience of the team at the CEGD. Based on the findings described above, the authors and collaborators acknowledged below designed the GQI to distinguish among four dimensions of genderqueer identity; (i) challenging the gender binary, (ii) social construction of gender, (iii) theoretical awareness, and (iv) fluidity over time. The present study reports the development and psychometric evaluation of the GQI scale to measure the prevalence of genderqueer identity among individuals referred to a gender identity clinic and community and online based samples of lesbian, gay, bisexual, and transgender individuals.

The goal of the larger research focus is to better conceptualize gender as a dynamic, latent construct that can be used in various clinical and community contexts across a broad spectrum of identities from mid-adolescence through late adulthood. Specifically, this paper reports on the primary efforts to refine and assess the reliability and validity of the GQI Scale by conducting a series of exploratory and confirmatory factor analyses across several samples. We expected to confirm four subscales, with one initially measured only among those pursuing medical transition and later refined to the broader population.

Methods

Study samples

Three different samples were used for the pilot exploratory factor analysis (EFA), two European samples, one from a Dutch gender clinic and one from Dutch lesbian, gay, and bisexual (LGB) community groups.

The third sample was a LGBTQ community sample recruited from two university lesbian, gay, bisexual, and transgender (LGBT) centers in the northwestern United States. The demographic variables collected for each sample varied slightly based on revisions to the study-related questionnaire, as it was piloted overtime in different locations. For example, the European samples were given the survey in Dutch and asked questions about sex assigned at birth (male, female, neither), gender identity at the moment (male, female, other), and self-identity (male, female, male and female, not male nor female, and I don't know); whereas, the combined U.S. LGBT community sample asked "what is your current sex" and "how would you describe your gender identity" with open ended response text boxes. An additional fourth sample was recruited via Amazon Mechanical (MTURK) with transgender and genderqueer participants to retest the factor structure and confirm the reliability after the final revisions. Specifically, the gender fluidity scale was broadened to incorporate all participants. The final revised items are included (see Tables 5 and 6).

The first sample is referred to as the Dutch Clinic sample consisting of 327 participants. All participants in the Dutch clinic sample were referred to the Center of Expertise on Gender Dysphoria at the VU University Medical Center in Amsterdam, the Netherlands for evaluation of gender dysphoria. Participants' ages ranged from 17 years to 68 years of age ($M = 31.36$; $SD = 12.49$). The majority of participants identified the sex they were assigned at birth as male (63.3%), whereas the remaining participants identified their assigned sex as female (36.7%).

The second sample is referred to as the Dutch Community sample and consisted of 290 LGB-identified individuals from LGB community groups in the Netherlands. Participants ages ranged from 18 years to 85 years of age ($M = 38.07$; $SD = 13.36$). The majority of participants identified their assigned sex as female (55.2%), whereas the remaining participants identified their assigned sex as male (44.8%).

The U.S. sample consisted of 150 individuals recruited from LGBT centers at two northwestern U.S. college campuses. Participants' ages ranged from 18 years to 61 years of age, 70% were under 30, ($M = 28.00$; $SD = 8.749$). Participants identified from a forced choice current sex as female (52.9%), male (31.9%), trans (1%), and intersex (<1%). In a separate open-ended text response, participants identified with a wide range of gender expressions, including: masculine (28.1%), feminine (42.3%), genderqueer/genderfluid (7.1%), transgender (1.4%), butch (3.8%), and agender (3.8%). A small percentage of participants chose not to identify their current sex or chose not to answer the question about gender expression.

The final, fourth, sample consisted of 510 lesbian, gay, bisexual, queer, transgender, and genderqueer identified participants recruited via an online survey forum. Inclusion criteria stipulated that participants must be between the ages of 18 and 74 and self-identified within the transgender or genderqueer populations. Participants were given the option to select an age range between 18 and 24 (11.4%), 25 and 34 (60.0%), 35 and 44 (20.6%), 45 and 54 (6.1%), 55 and 64 (1.8%), 65 and 74 (<1%). Just over half (55.1%) were assigned male at birth, while 43.9% were assigned female and <1% were unassigned. Current gender identities varied; 28.5% male, 28.3% female, 10.2% transfeminine, 7.7% transmasculine, 16.5% genderqueer, 3.5% non-binary, 4.7% genderfluid, and 1 participant each identified as genderfuck, agender, greygender, or undefined. A majority of participants identified as White (60.4%), although the sample was ethnically diverse, while 39.6% identified across a range of non-white identities (13.4% Black, 8.3% Latin 4.3% American Indian, 10.8%, Asian, 2.6% multiracial, <1% Middle Eastern).

Scale development

The GQI originally had 24 items and was developed in order to explore how people experience their gender as either male or female or neither female nor male. Development of measure items consisted of ongoing consultations between a faculty member in Family Social Science at the University of Minnesota—Twin Cities and clinical psychologists from the Center of Expertise on Gender Dysphoria of the VU University Medical Center. The creation of the GQI was based on theory, review of empirical research, qualitative interviews, analysis of empirical data, and clinical experience. Additionally, the iterative process to finalize and validate the measures has lasted several years and incorporated a number of researchers, trans, genderqueer, non-binary, and gender nonconforming LGB community members, clients at clinics, and clinicians from the U.S. and Europe.

We began by generating items that explored multiple dialectics of gender development and expression

such as a distinction of being born in the wrong body versus having a discomfort with the binary. We also wanted to measure peoples' experiences of gender as an expression of something that comes from within them, versus something they actively construct, something they are born with or something that develops. Do they want to explore and discuss gender, as an identity development process, or simply identify it and move on? After initial questions were developed, we began the iterative process of creating suitable instrument items and pilot testing.

Statistical analysis

To determine the factorial structure of the GQI an EFA was conducted first with responses from three samples: (1) European Clinical ($N = 327$); (2) European LGB Community ($N = 290$); (3) U.S. LGBT Community ($N = 150$) and (4) MTurk Trans and genderqueer sample ($N = 510$) (see Table 1). The EFA was conducted using SPSS 23 for Windows 7. Given the exploratory nature of the analyses and inconsistent missing data across items, cases were excluded pairwise to include as many members of the sample as possible. We allowed for oblimin (oblique) rotation of the constructs because we expected the factors to be correlated.

Researchers inspected the rotated factor matrix to determine item-factor loadings. Criteria for factor loadings included item values $\geq .40$ on the primary factor and $\leq .30$ on the other factors (Tabachnick, Fidell, & Osterlind, 2001). Based on the findings from the initial EFA some GQI items were slightly reworded. Items GQI 6, "It upsets me when people cannot tell if I am male or female" and GQI 17, "The way I think about my gender is something I talk with others about" were reconsidered and dropped before the CFA due to poor factor loadings in the EFA, and conceptual concerns with fit across a spectrum of gender identities.

The initial EFA was based on a 24-item GQI scale in which the gender fluidity items were intended only for people who sought medical services in support of physical transition. In the initial pilot of the GQI gender fluidity items were on a skip pattern and the participants in the European community and U.S. samples were not administered these items because the items were originally worded in such a way that distinctly asked about future plans to transition. The survey administered to the European LGB community sample and the U.S. LGBT campus

Table 1. Exploratory factor analysis of genderqueer identity (GQI) scale across four samples.

Variable	Dutch clinical trans	Dutch community LGB	US community LGBT	Online non-clinical LGBTQ
	N	N	N	N
	327	290	150	510
Challenge binary				
Genderqueer	.657	.782	.777	.778
Non-binary	.484	.592	.611	.688
Confuse	.706	.856	.810	.755
Bothfm	.677	.565	.539	.643
Unsure	.766	.843	.892	.832
Upsets	−.529	−.174 (dropped)	−.489	dropped
Factor	1st	3rd	3rd	1st
Eigenvalue	3.76	1.44	1.79	5.089
% of Variance	19.8	7.55	9.4	29.937
Cronbach's α	.739	.760	.804	.824
Social Construction				
Always ®	−.689	−.788	.660	.764
Natural ®	−.520	−.688	.543	.741
Time	.549	.702	−.657	−.429
Depends	.575	.728	−.709	−.481
Experiences	.580	.743	−.691	−.437
Talk	.183 (.355 on factor 2)	.568	−.531	dropped
Change	.624	.781	−.821	−.450
Factor	3rd	1st	2nd	3rd
Eigenvalue	1.7	6.56	2.10	1.28
% of Variance	8.95	34.53	11.06	7.55
Cronbach's α	.609	.852	.800	.670
Theoretical awareness				
Research	.512	.698	.651	.659
Convince	.767	.874	.861	.704
Expects	.738	.855	.810	.737
Free	.651	.752	.638	.664
Push	.573	.736	.626	.674
Open	.742	.864	.801	.746
Factor	2nd	2nd	1st	2nd
Eigenvalue	2.52	2.76	6.43	2.82
% of Variance	13.20	14.53	33.83	16.57
Cronbach's α	.767	.889	.843	.807

Note: Dropped indicates the item failed to load or loaded poorly and was dropped on subsequent versions.

sample did not receive the questions because they did not indicate a future desire to transition. The items were revised and shifted to be applicable regardless of clinical intention, opening the set of questions in the online survey with, "In the future…" A separate EFA was then run for the Gender Fluidity items among the European Clinical and the Online Community samples (see Table 2).

To determine the stability of the factorial structure of the GQI across groups, confirmatory multigroup analyses were performed using LISREL 8.8 for Windows (Jöreskog & Sörbom, 2006). Full information maximum likelihood minimization function was used in all analyses. Cross-validating the structure obtained with the EFA consisted of testing equality across groups

Table 2. Exploratory factor analysis of gender fluidity unidimensional model.

Variable	Dutch clinical transgender	Online non-clinical LGBTQ
	N	N
	327	510
Gender fluidity		
Traditional	.657	.628
No change	.484	.628
No question	.706	.661
Fluid®	−.677	−.473
Non-traditional®	−.766	−.609
Upsets	dropped	dropped
Eigenvalue	2.09	1.82
% of Variance	41.76	36.42
Cronbach's α	.645	.560

Note: Dropped indicates the item failed to load or loaded poorly.

of the factorial structure, the factor loadings, and the measurement errors for both the GQI three factor solution and the Gender Fluidity unidimensional model. The European clinical group served as the reference group for the all multigroup comparisons.

Results

Exploratory factor analysis

We performed an initial principal components factor analysis that determined that each subscale factored best into a three-factor solution, excluding the items from the *Gender Fluidity* subscale. For factor rotation we specified a three factor solution. The first factor consisted of 5-items and was labeled *Challenge the Binary*, the second factor consisted of seven items and was labeled *Social Construction*, and the third factor consisted of five items and was labeled *Theoretical Awareness*. Item loadings and total variance explained across all four samples are presented in Table 1. In almost all cases, each item has the highest factor loading on its intended scale with few exceptions. In all cases, the Cronbach's alpha for each scale was acceptable and was higher among the community samples and online sample than the clinical sample.

Unidimensional gender fluidity scale

In an initial EFA, the single factor solution for the Gender Fluidity subscale across the European Clinical and Online Non-clinical samples accounted for 41.76% and 48.28% of the variability, respectively (see Table 2). Five items loaded acceptably from the European Clinical sample, however, "In the future, I expect

my gender expression to be traditional" failed to load on the Online Non-clinical sample and "it upsets me when people cannot tell if I am male or female" failed to load across both samples, thus they were dropped from the CFA. Based on findings from the EFA, researchers confirmed the unidimensional Gender Fluidity subscale through confirmatory factor analysis. Given that the EFA specified a Gender Fluidity scale with five items for the clinical population and four items for the community population we know the scale is non-invariant across groups, thus a multigroup CFA was not performed.

Appropriate cutoff values were assessed according to Hu and Bentler's (1999) recommendations for good and adequate fit. Based on Hu and Bentler's (1999) recommendations, the expected values for a good model data fit is possible when the comparative fit index (CFI) index is above .95, Root Mean Square Error Approximation (RMSEA) index is below .06, and Standardized Root Mean Square Residual (SRMR) index is below .08. The unidimensional baseline CFA models indicated poor fit (CFI = .73 & .76, RMSEA = .19 & .27, SRMR = .082 & .11) for the European Clinical and online Community samples, respectively. Modification indices were employed and significantly improved the overall model fit for both samples (CFI = 1.00 & 1.00, RMSEA = .027 & .000, SRMR = .018 & .009), X^2difference (3) = 76.03, $p < .001$ for the clinical model and X^2difference(1) = 19.24, $p < .001$ for the community model.

Scale reliabilities for the modified CFA models were assessed using Hancock and Mueller's (2001) coefficient *H*, which provided statistical evidence whether items were consistent with each other. This method allows for more accurate scale reliability measurement when a model contains correlated measurement errors (Hancock & Mueller, 2001). Hancock and Mueller (2001) discussed that the magnitude of coefficient *H* for minimum reliability levels of .70 or higher are desirable. The scale reliability for the Gender Fluidity scales was .876 for the European Clinical sample and .877 for the online Community sample. Statistical significance was interpreted as a *z* score, ±1.96 with a critical value at an alpha level .05. All the standardized factor loadings were .30 or above, except "no question" was .23 in the MTurk sample. Finally, the average variance extracted was 31% for both samples. Overall, while the Gender Fluidity scale does operate differently in different social contexts, the confirmatory factor analysis has shown

that this unidimensional model of Gender Fluidity represents one viable relationship between the factor and the observed data across two different samples.

Three factor GQI multigroup confirmatory analysis

The multigroup confirmatory analyses were performed using only the Dutch Clinic and Dutch Community samples to asses group differences between persons referred to a clinical for medical transition and LGB community members. Preparation of the multigroup confirmatory analyses of the three-factor GQI scale consisted of fitting unidimensional models for each sample to provide a null hypothesis. The unidimensional models were rejected because all fit indices indicated poor model fit across both samples (CFI = .66 & .77, RMSEA = .15 & .23, SRMR = .11 & .14), resulting in significantly worse fit to the data than the three factor model, X^2difference(3) = 429.71, $p < .001$ for the clinical model and X^2difference(3) = 864.02, $p < .001$ for the community model.

We fit the theorized three-factor structure, which had adequate fit (CFI = .90 & .96, RMSEA = .068 & .077, SRMR = .064 & .076). Modification indices suggested adding correlated error variances between items "confuse" and "unsure"; "always" and "natural"; "depends" and "experiences"; and "convince" and "expects." Modification indices were applied due to a

strong theoretical underpinning that the items were closely correlated in wording and concept. The correlated error variances pointed to a possible measurement error due to similarity in item language and possibly translating the survey measure between the Dutch and English languages. For example, items, "The way I think about my gender was *always* this way" and "The way I think about my gender comes *naturally* from within me" are conceptually very similar questions though worded differently. After the implementation of modification indices model fit improved significantly (CFI = .95 & .97, RMSEA = .048 & .058, SRMR = .054 & .065), X^2difference(4) = 98.41, $p < .001$ for the clinical model and X^2difference(4) = 91.03, $p < .001$ for the community model.

The scale reliabilities for each factor across the two samples were as follows: Factor 1, Challenge the Binary ($H = .76$ & .82), Factor 2, Social Construction ($H = .80$ & .81), and Factor 3, Theoretical Awareness ($H = .79$ & .86). Based on Hancock and Mueller's recommendations, the three factor structure CFA within each sample proved to reliably measure what it purports to measure. Based on adequate model fit and scale reliability across the two samples, model evaluation proceeded to an inspection of the direction, magnitude, and significance of the parameter estimates. The results of the unstandardized parameter estimates, standard errors, communality, and average variance

Table 3. Parameter estimates from three-factor CFA model of genderqueer identity ($N = 327$; $N = 290$).

Model results	Estimates		S.E.		Est./S.E.		Communality (AVE)	
	Clin group	Com group	Clin group	Com group	Clin group	Com group	Clin group	Com group
CB By							(34.9%)	(43.8%)
Genderqueer	1.000	1.000			8.111	7.000	53.29%	51.84%
Non-binary	0.730	0.770	0.090	0.110	8.111	7.000	28.09%	21.16%
Confuse	0.500	0.700	0.070	0.070	7.143	10.000	26.01%	57.76%
Bothfm	0.760	0.900	0.090	0.100	8.444	9.000	40.96%	34.81%
Unsure	0.680	0.770	0.090	0.080	7.556	9.625	26.01%	53.29%
SC By							(21.6%)	(43.6%)
Always	1.000	1.000					19.36%	49.00%
Natural	0.430	0.420	0.090	0.060	4.778	7.000	9.00%	16.00%
Time	0.910	0.940	0.220	0.090	4.136	10.400	12.25%	47.61%
Depends	0.960	0.900	0.190	0.090	5.053	10.000	32.49%	42.25%
Experiences	1.130	0.920	0.230	0.100	4.913	9.200	27.04%	40.96%
Change	1.190	1.000	0.220	0.080	5.409	12.500	29.16%	65.61%
TA By							(34.9%)	(56.9%)
Research	1.000	1.000					15.21%	42.25%
Convince	1.960	1.220	0.340	0.100	5.765	12.200	40.96%	68.89%
Expects	1.880	1.160	0.330	0.080	5.697	14.500	38.44%	64.00%
Free	1.480	0.820	0.270	0.100	5.481	8.200	32.49%	43.56%
Push	1.430	1.010	0.260	0.100	5.500	10.100	30.25%	51.84%
Open	1.980	1.150	0.330	0.100	6.000	11.500	51.84%	70.56%

Note: Estimates, unstandardized parameter estimates; S.E, standard error; Est/S.E., test statistic (*z* value); AVE, average variance extracted for each factor and each group. Communality, the proportion of variance explained by the latent factor. Clin, is the clinical sample and Com is the community sample. The unstandardized factor loadings of genderqueer, always, and research are 1.0 because they are reference indicators, (Brown, 2015, p. 126).

extracted for each group are presented in Table 3. The results presented in Table 3 indicated that every freely estimated parameter was statistically significant. Results of the baseline CFA models for the Dutch clinical and community samples indicated good model fit, good scale reliability, statistically significant factor loadings that loaded properly on their latent structures, stable parameters, and good average extracted variance and variance explained.

All decisions from the single-group CFA were carried forward to the multigroup analyses. To test measurement invariance, this study used a step-down approach to multigroup analysis, thus researchers began with a fully constrained model and released parameters incrementally (Brown, 2015). The model fit for the constrained multigroup analysis indicated adequate to poor fit (CFI = .91, RMSEA = .081, SRMR = G1.10 & G2 .12). Modification indices were applied to achieve better model fit. One by one, we released constraints and evaluated overall model fit and areas of model strain.

The final, best fitting, model included freed factor loadings, freed error variances, freed factor covariance and freed error covariance (CFI = .97, RMSEA = .056, SRMR = G1.06 & G2 .08). We employed a X^2 difference test as a statistical comparative evaluation on the constrained solution against the freed solution X^2 difference(38) = 378.96, $p < .001$. Significant improvement of the model fit for the unconstrained solution indicated non-invariance and that the items have different relationships to the latent factor within different samples (see Table 4.).

Results from the confirmatory factor analyses have shown that the GQI consists of three dimensions and 17-items with responses recorded on a 5 point Likert scale. Construct validity was first assessed with a single-group confirmatory factor analysis in which factors were viewed to load significantly on their intended factor with loadings of ≥.40 and adequate reliability (Tabachnick et al., 2001). The individual CFA models showed good fit and showed that all the path coefficients were significant regardless of social context. Similar outcomes were found for the multigroup CFA. Hancock and Mueller's (2001) coefficient H for the three dimensions across the two groups, clinical and community, revealed that all of the H coefficients were adequate, ranging from .76 to .86 for the individual CFAs and .76 to .92 for the multigroup analyses. Results indicated acceptable scale reliability and internal consistency for interpreting genderqueer identity among individuals within clinical and community LGBTQ-identified samples.

Figures 1 and 2 show the standardized solutions for both Dutch clinical and community groups (respectively)

Table 4. Parameter estimates from three-factor multigroup CFA model of genderqueer identity ($N = 327$; $N = 290$).

Model results	Estimates		S.E.		Est./S.E.		Communality (AVE)	
	Clin group	Com group	Clin group	Com group	Clin group	Com group	Clin group	Com group
CB By							(35.1%)	(40.0%)
Genderqueer	1.000	1.000					51.84%	51.84%
Non-binary	0.770	0.740	0.090	0.100	8.556	7.400	25.00%	24.01%
Confuse	0.530	0.680	0.070	0.060	7.571	11.333	30.25%	49.00%
Bothfm	0.810	0.870	0.080	0.090	10.125	9.667	34.81%	40.96%
Unsure	0.720	0.740	0.090	0.070	8.000	10.571	33.64%	36.00%
SC By							(24.2%)	(39.4%)
Always	1.000	1.000					37.21%	37.21%
Natural	0.370	0.430	0.070	0.070	5.286	6.143	10.89%	14.44%
Time	0.750	1.050	0.130	0.100	5.769	10.500	19.36%	37.21%
Depends	0.640	1.010	0.090	0.100	7.111	10.100	21.16%	51.84%
Experiences	0.790	1.030	0.120	0.110	6.583	9.364	25.00%	40.96%
Change	0.850	1.120	0.120	0.090	7.083	12.444	31.36%	54.76%
TA By							(40.7%)	(49.0%)
Research	1.000	1.000					28.09%	28.09%
Convince	1.370	1.530	0.150	0.140	9.133	10.929	47.61%	59.29%
Expects	1.300	1.460	0.150	0.140	8.667	10.429	43.56%	54.76%
Free	1.030	1.030	0.120	0.110	8.583	9.364	37.21%	37.21%
Push	0.980	1.270	0.120	0.130	8.167	9.769	30.25%	50.41%
Open	1.380	1.450	0.140	0.130	9.857	11.154	57.76%	64.00%

Note: Estimates, unstandardized parameter estimates; S.E, standard error; Est/S.E., test statistic (*z* value); AVE, average variance extracted for each factor and each group. Communality, the proportion of variance explained by the latent factor. Clin, is the clinical sample and Com is the Dutch community sample. The unstandardized factor loadings of genderqueer, always, and research are 1.0 because they are reference indicators, subsequently, the standard errors of these estimate are 0.0 and thus no significance test is available (Brown, 2015, p. 126).
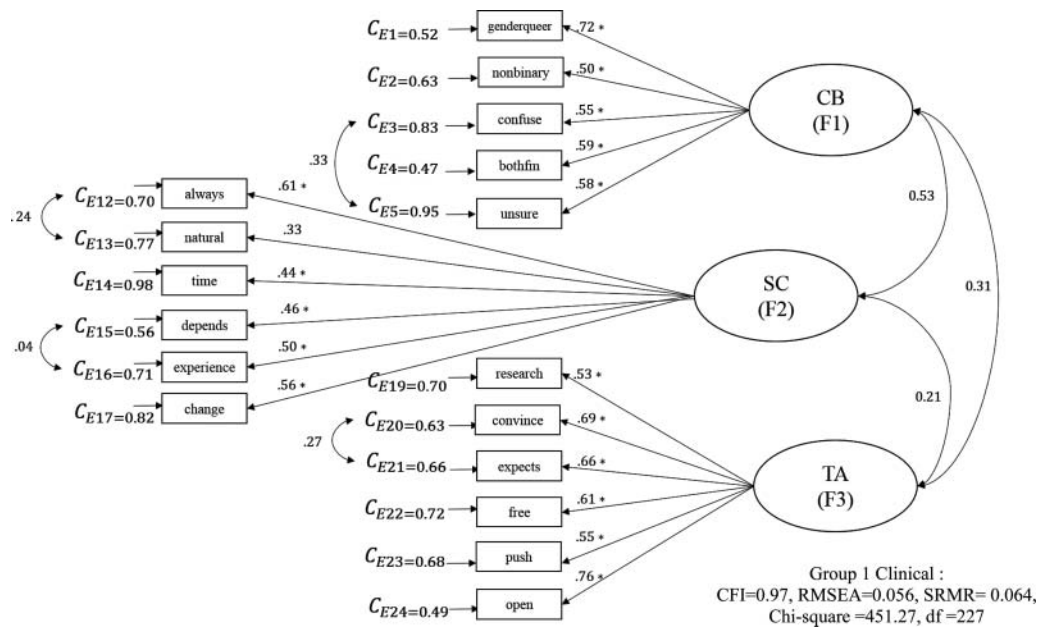
Figure 1. Multigroup Group 1 (Clinical) Three-Factor CFA standardized solution for the GQI (*N* = 327; 290).

within a multigroup analysis. All standardized factor loadings ranged from .38 to .90 and all freely estimated factor loadings were statistically significant at $p < .05$, and salient across the two samples. Modification indices pointed to some areas of strain, however, because the overall model fit was good and the modifications did not make theoretical sense, they were not utilized. All other indicators provided evidence that the model was working well across groups. Simultaneous analysis of the two groups revealed that the null hypotheses of invariance across groups and

factorial invariance were rejected. Thus, results indicated that although a similar three-factor solution provided the best fit across clinical and community samples, the underlying concepts were understood differently between the groups.

Discussion

The aim of the present study was to develop and validate a dynamic measure of genderqueer identity that
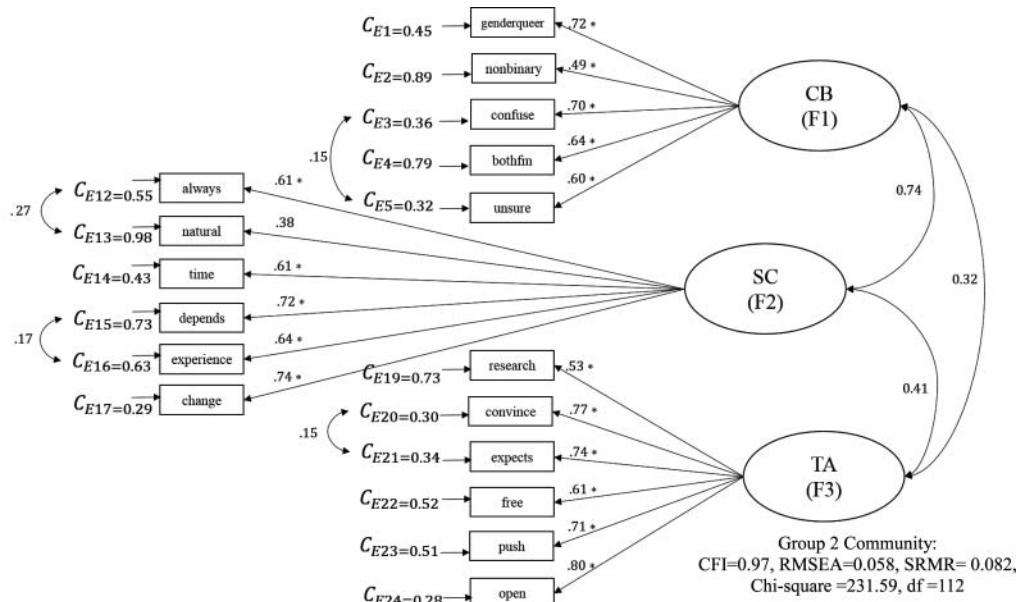
Note: All coefficients with an * are significant at p < .001. F1: Challenge Binary, F2: Social Construction, F3: Theoretical Awareness

Figure 2. Multigroup Group 2 (Community) Three-Factor CFA standardized solution for the GQI (*N* = 327; 290).

can be used in various clinical and community contexts, with a variety of gender identities and expressions, across a good portion of the lifespan. The development of the GQI fills critical gaps in gender-related measurement including the ability to assess multiple dimensions of gender identity. The GQI instrument was developed as a result of extensive literature review, the qualitative and clinical research of experts and collaborators in multiple continents, and a pilot study across three groups resulting in an EFA. The EFA exposed potential kinks in the original 24-item GQI measure, which were refined and retested. Researchers evaluated and cross-validated the hypothesized factor structure and determined that the three factor GQI subscales and the unidimensional Gender Fluidity measure yielded internally consistent and valid scores among transgender individuals seeking clinical treatment and LGB individuals within a community setting. Thus, the final GQI with subscales (see Tables 5 and 6) may serve as a valuable measurement tool to examine the multiple domains of gender identity and capture variability in genderqueer identity across a broad spectrum of physical expression, personal identification, and belief systems.

As theorized, the first factor, Challenge the Binary, consisted of five items that focused on the extent to which personal expressions of gender identity were not accounted for within the gender binary. This expression of genderqueer identity examines how a person may show identity and express elements of

Table 5. Finalized genderqueer identity (GQI) scale questions and Likert scale.

Question items	Likert scale				
Subscale 1. Challenging the Binary	Strongly Disagree 0	Disagree 1	Neutral 2	Agree 3	Strongly Agree 4
The statements below are about your gender identity and expression. Please indicate to what degree you agree with each statement. 1. I am non−binary, genderqueer, or an identity other than male or female. 2. I don't want to be seen in the gender binary (as either male or female) 3. I try to deliberately confuse people about whether I am male or female. 4. I try to do things that are masculine and feminine at the same time. 5. I enjoy it when people are not sure if I am male or female.					
Subscale 2. Social Construction	Strongly Disagree 0	Disagree 1	Neutral 2	Agree 3	Strongly Agree 4
The statements below are about how you understand your gender. Please indicate to what degree you agree with each statement. 6. The way I think about my gender has always been the same. [R] 7. My gender comes naturally from within me. [R] 8. My gender is something I have spent a lot of time figuring out. 9. The way I show my gender changes depending on who I am with. 10. The way I think about my gender has been influenced by experiences in my life. 11. The way I think about my gender will probably continue to change further as I age. *. I talk a lot with others about gender.					
Subscale 3. Theoretical Awareness	Strongly Disagree 0	Disagree 1	Neutral 2	Agree 3	Strongly Agree 4
The statements below are about your political and theoretical awareness of gender. Please indicate to what degree you agree with each statement. 12. I have done research about gender theory and gender roles. 13. I try to convince others that society should not insist on a gender binary. 14. I try to convince others that society expects people to be too gender conforming. 15. Around me, I make sure people are free to express whatever gender roles they want. 16. The way I show my gender is important because I push society to question traditional gender roles. 17. I encourage others to be more open minded about gender and gender roles.					

Note: Subscales were developed based on prior qualitative analyses and research literature review of gender identity.
*Item loads poorly in the EFA clinical sample, and inconsistently in the CFA non-clinical sample. Optional to include if needed for other purposes.

Table 6. Finalized gender fluidity subscale survey questions and Likert scale.

Question Items	Likert Scale				
Subscale 4. Gender Fluidity	Strongly Disagree 0	Disagree 1	Neutral 2	Agree 3	Strongly Agree 4

The statements below are about how fluid you think your gender will be in the future.
  Please indicate to what degree you agree each statement following the stem prompt.
1. In the future, my gender expression will be traditional. [®]
2. In the future, it will upset me if people misgender me.
3. The way I show my gender will probably be mostly the same from day to day. [®]
4. In the future, I expect that people will rarely question my gender. [®]
5. In the future, I think my gender will be fluid or change over time.
6. I will have a non-traditional gender role (be gender non-conforming).

Note: Gender Fluidity items were originally drafted with a skip patter for people seeking medical transition. Ultimately the items were modified and tested on a broader sample, but needed to be factored separately because of the wording changes. Future iterations could factor *Fluidity* with the other subscales, or independently as needed. Items marked with [®] should be reverse scored.

both masculine and feminine identity in everyday life. A person scoring high on challenging the binary, may not necessarily score high on the other subscales, suggesting that expression may or may not feel socially constructed, theoretically grounded, or variable from day to day. Trans persons do not necessarily challenge the belief that gender is a dichotomous binary with two contrasting identity categories representing masculinity on one end and femininity on the other (Bauer et al., 2009; Diamond et al., 2011; Serano, 2007). Challenging the binary, thus, can include both the transgressive element of undoing, resisting, or protesting the idea that gender identity is discrete and dichotomous, and also an expression of gender that exists between or outside the binary, regardless of transgressive intent. Genderqueer persons, and those who blur the lines and boundaries between the relationship of assigned sex at birth and affirmed gender identities and expressions represent a more expansive definition of the variable ways in which gender identity is subjectively experienced and expressed (Beek et al., 2015).

The second factor, Social Construction, consisted of six items that focused on the extent to which persons felt their gender was socially constructed. This taps into interpretations about the aspects of gender as performance, and as something that develops with intentional influence from the context within which one lives. In this construct, an individual may score highly without a physical expression of genderqueer identity, making it a useful measure across a broad range of gender identities. Additionally, some research participants described genderqueer expressions that felt more innate, while others described deliberate efforts to construct an expression. This subscale assesses that

distinction. One item "My gender identity is something I talk with others about a lot" functioned better in community LGB samples than clinical trans samples. For this reason, it is listed as optional on Table 5. In the context of clinical assessment for gender dysphoria, there is a ceiling effect on the concept of talking with others about gender identity, inhibiting variability, and scale performance.

The third factor, Theoretical Awareness, consisted of six items that focused on the extent to which persons felt their gender identity was based on a political or theoretical understanding of gender. This subscale also complements the expression of genderqueer identity but does not require it. An individual may or may not have theoretical interest or study in gender theory as an element of gender identity, however, a notable subset of prior research participants and community members had described this, and it was supported in the literature review. The measure showed validity and contributed depth to the overall construct of genderqueer identity, as something richer and deeper than non-binary expression.

The first piloted version of the GQI was administered with the Gender Fluidity subscale on a skip pattern, such that, only individuals who indicated a desire to pursue steps toward medical transition were offered the questions. A distinct factor analysis was run for those items among the European clinic sample only. The single factor solution accounted for 41% of the variability, and five of the original six items loaded acceptably onto that factor. The item, "It upsets me when people cannot tell if I am male or female" did not function well on scales and was dropped. It tapped two issues: whether one is ever misgendered—using a word or pronoun for a person that does not reflect

their affirmed gender identity—and whether this is upsetting. A "strongly disagree" was too hard to interpret. The conceptual concern remains however, that it is worthwhile to know, especially in a clinical context, whether or not an individual is upset (or conversely enjoys) the experience of being misgendered.

Based on the EFA, we decided to recreate the Gender Fluidity subscale so that it would be salient regardless of a desire for medical transition. The change was subtle, altering the stem prompt from "once I have completed my transition" to "in the future." The EFA provided adequate evidence to support that Gender Fluidity was operating in clinical and community samples as intended. However, it still functioned better when factored separately because it had been measured slightly differently in the two samples. Future studies will need to further refine the concept of fluid gender presentation and its measurement across a spectrum of gender identity. It is worth noting here that in order to be fluid, change across time, one must necessarily deviate from the binary somewhat at least sometimes. Conversely, to deviate from the binary, it is not necessary at all to be fluid, change from day to day. One can be quite genderqueer in expression, yet fundamentally non-fluid, that is the same from day to day.

Implications

Our long-term goal is to better understand the development of gender identity. We focus on gender identity development processes (e.g., renegotiation, emergent properties, interchange between the individual and the environment) to begin to measure variability among non-binary individuals, those who hold a middle ground either between or outside the male-female gender binary. The objective was to validate a measure that could be used to locate within group variation and nuanced gender identity. The development of the GQI scale has clinical and public health implications, in that, the ways in which non-binary identity develops may be associated with different gender affirmative treatment seeking decisions within and across different cultures. The process of non-binary identity development may vary between different cultures, and may be differentially associated with treatment seeking behaviors based on contextual influences. This study has validated an instrument to measure gender on a spectrum rather than as a categorization of transgender or not. With spectrum-based measurement, we can better explore how to support healthy development in a broad range of gender diverse people.

Limitations and future directions

There are limitations in generalizability due to the age range and ethnic distribution of the samples. Two samples were mostly ethnically European Caucasian, another was collected online from college campuses that were 85% white, and the final one was more ethnically diverse (50% non-white). All of the samples were collected on adults, with somewhat of a trend toward young adults in the 20s and 30s. The online samples limit the respondent availability and yield less reliable data given the anonymous nature of an online sample. Future research should repeat psychometric work with different clinical- and community-based samples in additional cultural contexts such as U.S. clinics, additional communities from other world regions, more ethnically diverse communities, and with adolescents and populations over age 50. Also, Diamond and Butterworth (2008) have argued that gender identity is fluid and that gender identity may fluctuate over time, suggesting repeated measures longitudinal studies would be a useful next step. The GQI may function differently with a younger or older population based on discrete developmental stages. We have tested this measure across two distinct cultural contexts with very different approaches to health care and coverage of medical costs related to transition. The instrument functions well in both cultures, and provides evidence of measurable genderqueer identity characteristics in both environments. This study should be replicated across more ethnically and developmentally diverse populations in order to further confirm and validate the structure, its relationship to the observed indicators, and the multiple subscales.

Acknowledgments

Declaration of conflict of interest

The authors declare they have no conflicts of interest.

References

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

Beek, T. F., Kreukels, B. P. C., Cohen-Kettenis, P. T., & Steensma, T. D. (2015). Partial treatment requests and underlying motives of applicants for gender affirming interventions. *Journal of Sexual Medicine*, *12*(11), 2201–2205. doi:10.1111/jsm.13033

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162. doi:10.1037/h0036215

Bockting, W. O. (2014). Transgender identity development. In D. L. Tolman & L. Diamond (Eds.), *APA handbook of sexuality and psychology, Vol. 1: Person-based approaches.* (pp. 739–758). Washington, DC: American Psychological Association.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Publications.

Budge, S. L., Katz-Wise, S. L., Tebbe, E. N., Howard, K. A. S., Schneider, C. L., & Rodriguez, A. (2013). Transgender emotional and coping processes. *The Counseling Psychologist*, *41*(4), 601–647. doi:10.1177/0011000011432753

Butler, J. (1988). Performative acts and gender constitution: An essay in phenomenology and feminist theory. *Theatre Journal*, *40*(4), 519–531. doi:10.2307/3207893

Butler, J. (1990). *Gender trouble*. New York, NY: Routledge.

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 & Adolescent Psychiatry*, *36*(2), 263–271. doi:10.1097/00004583-199702000-00017

Deogracias, J. J., Johnson, L. L., Meyer-Bahlburg, H. F., 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*(4), 370–379. doi:10.1080/00224490701586730

Diamond, L. M., & Butterworth, M. R. (2008). Questioning gender and sexual identity: Dynamic links over time. *Sex Roles*, *59*(5-6), 365–376. doi:10.1007/s11199-008-9425-3

Diamond, L. M., Pardo, S. T., & Butterworth, M. R. (2011). Transgender experience and identity. In S. J. Schwartz, K. Luyckx, & V. L. Vignoles (Eds.), *Handbook of identity theory and research* (pp. 629–647). New York, NY: Springer.

dickey, l. m., Hendricks, M. L., & Bockting, W. O. (2016). Innovations in research with transgender and gender nonconforming people and their communities. *Psychology of Sexual Orientation and Gender Diversity*, *3*(2), 187–194. doi:10.1037/sgd0000158

Doan, P. L. (2016). To count or not to count: Queering measurement and the transgender community. *WSQ: Women's Studies Quarterly*, *44*(3-4), 89–110. doi:10.1353/wsq.2016.0037

Egan, S., & Perry, D. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, *37*(4), 451–463. doi:10.1037/0012-1649.37.4.451

Fassinger, R. E., & Arseneau, J. R. (2007). "I'd rather get wet than be under that umbrella": Differentiating the experiences and identities of lesbian, gay, bisexual, and transgender people. In K. J., Bieschke, R. M., Perez, & K. A. DeBord (Eds.), *Handbook of counseling and psychotherapy with lesbian, gay, bisexual, and transgender clients (2nd ed.).* (2nd ed., pp. 19–49). Washington, DC: American Psychological Association. doi:10.1037/11482-001

Grant, J., Mottet, L., Tanis, J., Harrison, J., Herman, J., & Keisling, M. (2011). *Injustice at every turn: A report of National Transgender Discrimination Survey*. Washington, DC: The National Gay and Lesbian Task Force and the National Center for Transgender Equality. doi:10.1016/S0016-7878(90)80026-2

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit, & D. S Sörbom (Eds.), Structural equation modeling: Present and future a festschrift in honor of Karl Jöreskog (pp. 195-216). Lincolnwood, IL: Scientific Software International.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit, & D. S Sörbom (Eds.), *Structural Equation Modeling: Present and Future, a festschrift in honor of Karl Jöreskog.* (pp. 195-216). Lincolnwood, IL: Scientific Software International.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. doi:10.1080/10705519909540118

Harrison, J., Grant, J., & Herman, J. L. (2012). A gender not listed here : Genderqueers, gender rebels, and otherwise in the National Transgender Discrimination Survey. *LGBTQ Policy Journal at the Harvard Kennedy School*, *2*, 13–24.

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. doi:10.1177/0361684312442161

Johnson, A. H. (2016). Transnormativity: A new concept and its validation through documentary film about transgender men. *Sociological Inquiry*, *86*(4), 465–491. doi:10.1111/soin.12127

Jöreskog, K. G., & Sörbom, D. (2006). *LISREL 8.80* [Software]. Chicago, IL: Scientific Software International.

McGuire, J. K., Kuvalanka, K. A., Catalpa, J. M., & Toomey, R. B. (2016). Transfamily theory: How the presence of trans[*] family members informs gender development in families.

Journal of Family Theory & Review, 8(1), 60–73. doi:10.1111/jftr.12125

Money, J. (1975). Ablatio penis: Normal male infant sex-reassigned as a girl. *Archives of Sexual Behavior*, 4(1), 65–671. doi:10.1007/BF01541887

Nagoshi, J. L., & Brzuzy, S. (2010). Transgender theory: Embodying research and practice. *Affilia*, 25(4), 431–443. doi:10.1177/0886109910384068

Nagoshi, J. L., Brzuzy, S., & Terrell, H. K. (2012). Deconstructing the complex perceptions of gender roles, gender identity, and sexual orientation among transgender individuals. *Feminism & Psychology*, 22(4), 405–422. doi:10.1177/0959353512461929

Olson, K. R., Durwood, L., DeMeules, M., & McLaughlin, K. A. (2016). Mental health of transgender children who are supported in their identities. *Pediatrics*, 137(3), 1–8. doi:10.1542/peds.2015-3223

Pollock, L., & Eyre, S. L. (2012). Growth into manhood: Identity development among female-to-male transgender youth. *Culture, Health & Sexuality*, 14(2), 209–222. doi:10.1080/13691058.2011.636072

Rankin, S. S., & Garvey, J. C. (2015). Identifying, quantifying, and operationalizing queer−spectrum and trans−spectrum students: Assessment and research in student affairs. *New Directions for Student Services*, 2015(152), 73–84. doi:10.1002/ss.20146

Reisner, S. L., Conron, K. J., Scout, Baker, K., Herman, J. L., Lombardi, E., … Matthews, A. K. (2015). "Counting" transgender and gender-nonconforming adults in health research: Recommendations from the gender identity in US surveillance group. *Transgender Studies Quarterly*, 2(1), 34–57. doi:10.1215/23289252-2848877

Richards, C., Bouman, W. P., & Barker, M-J. (2017). *Genderqueer and non-binary genders*. London, UK: Palgrave Macmillan.

Richards, C., Bouman, W. P., Seal, L., Barker, M-J., Nieder, T. O., & T'Sjoen, G. (2016). Non-binary or genderqueer genders. *International Review of Psychiatry*, 28(1), 95–102. doi:10.3109/09540261.2015.1106446

Rosario, M., Schrimshaw, E. W., Hunter, J., & Levy-Warren, A. (2009). The coming-out process of young lesbian and bisexual women: Are there butch/femme differences in sexual identity development? *Archives of Sexual Behavior*, 38(1), 34–49. doi:10.1007/s10508-007-9221-0

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. doi:10.1007/s10508-016-0702-x

Serano, J. (2007). *Whipping Girl: A transsexual women on sexism and the scapegoating of femininity*. Berekley, CA: Seal Press.

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 the gender identity/gender dysphoria questionnaire for adolescents and adults: further validity evidence. *The Journal of Sex Research*, 47(1), 49–58. doi:10.1080/00224490902898728

Stryker, S. (2008). *Transgender history*. Berekley, CA: Seal Press.

Tabachnick, B. G., Fidell, L. S., & Osterlind, S. J. (2001). *Using multivariate statistics* (4th ed.). New York, NY: Harper Collins.

West, C., & Zimmerman, D. H. (1987). Doing gender. *Gender & Society*, 1(2), 125–151. doi:10.1177/0891243287001002002

Wilchins, R. A. (2004). *Queer theory, gender theory: An instant primer*. Los Angeles, CA: Alyson Books.

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), 469–483. doi:10.1007/s11199-006-9019-x