

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Trans Collaborations Academic Papers

Psychology, Department of

2021

Psychometric Evaluation of the Transgender Congruence Scale

T. Zachary Huit

Allura L. Ralston

J, Kyle Haws

Natalie Holt

Debra A. Hope

See next page for additional authors

Follow this and additional works at: <https://digitalcommons.unl.edu/trans>



Part of the [Counseling Commons](#), [Developmental Psychology Commons](#), [Development Studies Commons](#), [Gender and Sexuality Commons](#), [Mental and Social Health Commons](#), [Other Psychiatry and Psychology Commons](#), [Other Psychology Commons](#), and the [Other Sociology Commons](#)

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Trans Collaborations Academic Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors

T. Zachary Huit; Allura L. Ralston; J. Kyle Haws; Natalie Holt; Debra A. Hope; Jae Puckett; Richard MocarSKI; and Nathan Woodruff

Psychometric Evaluation of the Transgender Congruence Scale

T. Zachary Huit,¹ Allura L. Ralston,¹ J. Kyle Haws,¹
Natalie R. Holt,¹ Debra A. Hope,¹ Jae A. Puckett,²
Richard A. MocarSKI,³ and Nathan Woodruff⁴

1 Department of Psychology, University of Nebraska–Lincoln, 238 Burnett Hall,
Lincoln, NE 68588-0308, USA

2 Department of Psychology, Michigan State University, East Lansing, MI, USA

3 Department of Communication, University of Nebraska at Kearney, Kearney,
NE, USA

4 Trans Collaborations Local Community Board, Lincoln, NE, USA

Correspondence – T. Zachary Huit, zach.huit@gmail.com

Abstract

Introduction Despite increased attention to transgender and gender diverse (TGD) issues in psychological literature during the past decade, gaps remain for psychometric validation of TGD-specific measures. Kozee et al. (*Psychology of Women Quarterly* 36(2):179–196, 2012) addressed such gaps by creating the Transgender Congruence Scale (TCS), measuring gender acceptance and feelings of gender congruence between internal and external attributes across a broad range of gender identities. The current study extended Kozee and colleagues' work by further examining the psychometric properties of the TCS.

Methods Between October and November of 2017, 210 transmasculine, transfeminine, and gender diverse adults between ages 19 and 73 completed online surveys

Published in *Sexuality Research and Social Policy* (2021)

doi:10.1007/s13178-021-00659-7

Copyright © 2021 by the authors; published by Springer Science+Business Media, LLC.

Used by permission.

Accepted for publication 12 October 2021.

containing demographic, gender identity-specific, and well-being measures including the TCS.

Results Results of confirmatory factor analysis replicated the original two-factor model (*Appearance Congruence* and *Gender Identity Acceptance*), with the elimination of two poorly loading items, resulting in a 10-item reduced model (*TCS-10*). Higher scores on *TCS-10* were associated with positive scores on gender-related well-being, congruence, and pride, positive affect, and life satisfaction, as well as lower scores on gender-related dysphoria, non-affirmation, internalized transphobia, and marginalization. There were modest but significant associations between *Gender Identity Congruence* and both depression and negative affect.

Conclusions Despite limitations of sample size and diversity of identities, the study reaffirmed the utility of the TCS as an overall construct of gender identity congruence with a 10-item reduced structure related to other established TGD constructs.

Policy Implications Though socio-political climate is the ultimate domain for alleviating TGD stigma and discrimination, factors such as gender congruence are essential areas of focus to foster resiliency.

Keywords: Transgender, Gender diverse, Identity formation, Psychometrics, Measurement, Well-being

Transgender and gender diverse¹ (TGD) communities have historically been underrepresented in research contexts, often being combined with samples of sexual minority participants, masking any distinct findings for TGD participants. However, in more recent years, TGD people have received growing attention in empirical studies. A Google Scholar search on October 6, 2020, specifying “Transgender” in the title, yielded over 3,630 results between the years of 2011 and 2014, almost tripling in number between 2015 and 2020 with over 11,900 results. Despite this increase in attention, there remain areas for further growth and salient gaps in knowledge, including recognizing the diversities of gender identities and a lack of validated assessment instruments for key TGD-related constructs.

Research has often perpetuated the exclusion of many TGD people by failing to capture gender outside of the male/ female binary, or by forcing people to select “other” if the binary designation does not

¹ Although some of the cited literature in this paper uses the term non-binary or gender non-conforming, at the recommendation of our Local Community Board, we will refer to those identifying as a gender different from their sex assigned at birth, but outside of male-female dichotomies, as *gender diverse*.

fit. An additional critique of TGD-related studies is the overemphasis on gender transition, furthering the misperception that all TGD individuals wish to pursue gender affirmation procedures and failing to recognize the individual nature of TGD identities and experiences (Benson, 2013; Puckett et al., 2018). Given that TGD communities are heterogeneous and have many intersectional identities and experiences (de Vries, 2012; Kuper et al., 2012), assumptions of gender binary and desires for transition perpetuate stigma and exclude a range of TGD people (Chang & Chung, 2015). In the current study, TGD refers to a broad range of gender diversity, excluding cisgender people, and including anyone identifying as a gender that differs from their sex assigned at birth.

Another salient gap in TGD literature is the lack of validated and reliable measurements of TGD-specific constructs. In their review, Shulman et al. (2017) identified eight TGD-specific measures, noting that even those measures had limited evidence of their psychometric properties outside of the original publication. Furthermore, many measures used with TGD individuals adapt previously established constructs and measures rather than starting with TGD experiences. For example, the *Transgender Stigma Scale* (Mizock & Mueser, 2014) was adapted from a measure of mental illness stigma (King et al., 2007). While such an approach serves as a short-term practical solution to the dearth of TGD-specific measures, the experiences of TGD individuals may not be fully captured by these adaptations (Hope et al., 2016). Rather, what is needed is specific measurement development and validation with TGD samples.

Kozee et al. (2012) addressed these major issues by creating the *Transgender Congruence Scale* (TCS) to assess how TGD people evaluate their personal acceptance of their gender and how closely their gender expression aligns with their identity. They defined transgender congruence as “the degree to which transgender individuals feel genuine, authentic, and comfortable within their external appearance/presence and accept their genuine identity rather than the socially prescribed identity” (p. 181). The constructs that they cite as important are the following: (a) clarity in identifying their gender (“Self-Image”); (b) how congruent they perceive their internal and external representations of gender (“Self-Reflection”); and (c) ability to express gender (“Self-Expression”). Kozee et al. (2012) indicated that

these constructs were designed to be more inclusive of gender diversity outside of binary identities and experiences.

The initial validation study of the TCS was conducted in two phases by Kozee et al. (2012). In the first phase, a team of academic researchers developed and refined items. These items were rated by four TGD individuals who deemed the 15 retained items as indicative of the desired content and sufficient in clarity. A sample of 162 TGD people between the ages of 18 and 75 completed a battery of assessments including the 15-item TCS and other measures of life meaning and satisfaction, body satisfaction, anxiety, depression, social desirability, and an inventory to assess the participants' steps in aligning their gender expression with their gender identity. Exploratory factor analysis yielded a final 14-item scale, with two factors, named *Appearance Congruence* (external appearance) and *Gender Identity Acceptance* (self-acceptance). This structure was further examined in the second phase of the validation using confirmatory factor analysis (CFA) among a sample of 342 TGD participants between the ages of 18 and 72. The results of the second study indicated a reduction of items to a 12-item scale, while retaining the original two-factor solution. Additionally, they found higher scores on these factors related to greater life meaning and satisfaction, as well as lower anxiety and depression. Other research has also supported the construct validity of the TCS (Chodzen et al., 2019; Jackman et al., 2018; Jones et al., 2019a; McLemore, 2015; van den Brink et al., 2019). TCS scores are negatively related to frequency of misgendering experiences, depression, negative affect, nonsuicidal self-injury, and rumination about gender identity; conversely, TCS scores are positively related to self-esteem (appearance and social), gender congruence, and life satisfaction, as well as overall quality of life. Internal consistency for the TCS tends to be good (Cronbach's α ranging from 0.79 to 0.96).

Although this growing body of work has addressed important gaps in TGD research related to gender diversity and helped rectify the dearth of empirically supported constructs that can be reliably assessed, some important gaps remain. First, Kozee et al. (2012) desired to capture a broader range of gender diversity outside of binary constructs (e.g., transgender man, transgender woman), but there was a limited proportion of individuals in their initial samples who strictly identified as gender diverse, without also identifying in a transgender

category (only 6 out of 162 for study 1 and 22 of 342 for study 2). While they found no significant differences in factor structure between those identifying as binary and gender diverse, limited group size may have dampened some of those differences. Jones et al. (2019b) similarly found no differences in *Appearance Congruence* between gender diverse and binary transgender participants; however, this was among a larger cisgender sample with a smaller proportion of transgender and gender diverse participants. Second, the employment of the TCS has varied regarding use of only the *Appearance Congruence* subscale, both subscales, or only the total TCS score, making direct comparisons more challenging. Finally, despite the broad implementation of the TCS, no other studies to date have examined other psychometric properties such as factor structure of the scale since the original validation study (Kozee et al., 2012).

As such, the purpose of the current study was to replicate and extend the original psychometric evaluation of the TCS using confirmatory factor analysis with an additional sample of TGD people, including a larger proportion of individuals identifying with genders outside of the binary. Since Kozee et al. (2012), a number of validated TGD-specific measures have been published and prominent general measures of mental health and well-being have been used extensively with TGD samples to support their validity. This provides new opportunities to further evaluate the psychometric properties of the TCS that were not available to Kozee et al. (2012), building on their groundbreaking work.

Overall, we hypothesized that we would replicate and extend the original validation study for the TCS. Specifically, we hypothesized that the *TCS would be internally consistent*, and the *original two-factor solution would be replicated* in the new data set. Also, we hypothesized that *greater gender congruence* (total score, appearance congruence, and gender identity acceptance) would be *associated with higher well-being on TGD-specific measures* (i.e., higher scores on gender-related functioning and gender pride; lower scores on gender-related dysphoria, discrimination, rejection, victimization, non-affirmation, internalized transphobia, negative expectations for the future, and nondisclosure of gender identity), and *better overall well-being* (i.e., lower ratings of anxiety, depression, and negative affect; greater ratings of positive affect and life satisfaction). Lastly, we expected that

while there may be slight differences for TCS scores between binary and diverse gender categories, there would be *no significant mean differences in TCS scores between those who identify as transfeminine, transmasculine, and gender diverse.*

Methods

Participants

Two hundred and ten TGD individuals completed a battery of on-line questionnaires, including the TCS. Participants were between 19 and 73 years old ($M = 29.93$, $SD = 9.42$). Sixty-one people (29%) had unknown ages; however, they confirmed that they were over the age of 19 per study inclusion criteria (with 19 being Nebraska's legal age of majority). Participants described their gender identity via an open response box as well as selecting one of three set categories. For the set categories, 74 (35.2%) participants identified as transwoman/trans woman/MTF (male-to-female)/woman, 75 participants (35.7%) as transman/trans man/FTM (female-to-male)/man, and 46 (21.9%) as nonbinary/gender nonconforming/ genderqueer/agender/bigender/another gender minority in the fixed categories. For brevity, these groups will be referred to as *transfeminine*, *transmasculine*, and *gender diverse*, respectively. Fifteen (7.1%) participants did not choose from the fixed options. Additionally, the free response field yielded a broad range of identities that formed too many categories for meaningful statistical analysis. There was a diverse representation of sexual orientation among participants: 95 (45.2%) identifying as heterosexual, 32 (15.2%) as gay, 21 (10%) as bisexual, 16 (7.6%) as lesbian, 15 (7.1%) as queer-identified, 6 (2.9%) as pansexual, and 4 participants (1.9%) as asexual. Five (2.4%) described their sexual orientation in a different manner, and 16 (7.6%) did not respond to sexual orientation.

A majority of respondents identified as European American/ Caucasian/White ($n = 157$, 74.8%). Twelve (5.7%) identified as African American/Black, 11 (5.2%) as Native American/American Indian/Alaskan Native, 6 (2.9%) as Asian American/Pacific Islander, 9 (4.3%) as Latino/a/x, and 9 (4.3%) as Hispanic. Eleven (5.2%) identified as

more than one race/ethnicity. Two (1%) opted to describe their racial/ ethnic identity in another way, and 16 (7.6%) did not respond to racial/ethnic identity.

Procedures

Participants were recruited via emails to listservs for LGBTQIA+ national organizations in the USA, as well as social media postings. Participants were directed to an anonymous link to a Qualtrics survey, with all data collection procedures occurring between October and November of 2017. TGD-identified participants over the age of 19 were asked to participate in the online study with a compensation of \$10. Those who accessed the survey link confirmed that they were over the age of 19, identified as TGD, and consented to participation before they were able to access the survey.

Individuals completed questionnaires regarding demographic information, general mental health, gender-related constructs, and the TCS. Order of administration of measures was randomized to reduce order effects. Participants also completed two additional measures that are not a part of the current analysis (Obasi, 2016). The data in the current analyses were collected as part of another study that was aimed to explore the psychometric properties of the Trans Collaborations Clinical Check-in (TC³; Holt et al., 2019). Those who completed the survey were electronically delivered a \$10 online gift card within 24 hours after completing the survey if they provided an email address to receive the gift card (all email addresses were deleted following payment).

The flow of participant self-selection and exclusion is detailed in **Fig. 1**. Qualtrics metadata and demographic information were examined for duplicate responses, and only those who did not complete the TCS were excluded from the data set. There was no indication of duplicate responses after the removal of those who did not complete any of the included measures in the survey, in addition to those who did not complete the TCS items specifically. The remainder of participants retained for analysis were determined to be valid, individual responses. The mean completion time of the participants retained for analysis was 115 min (*median* = 23.78, *SD* = 511.75). There is a wide variability in completion time due to inclusion of participants that

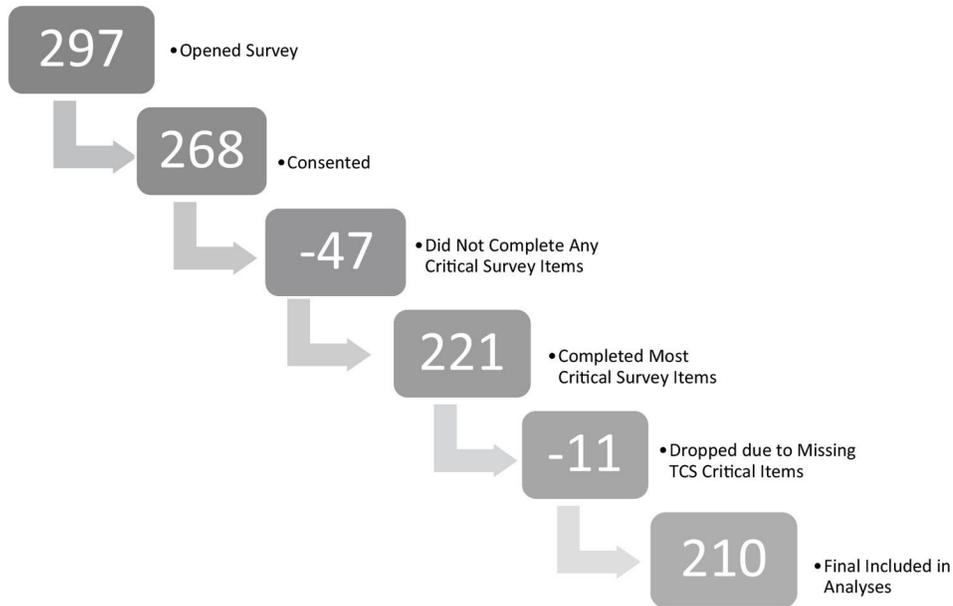


Fig. 1 Flow of participants through recruitment to data analyses

partially completed the survey battery but left it open, as the exclusion criterion in the current study was for non-completion of the TCS specifically. Among the various lengths for survey completion, there was adequate completion for most of the measures used in the current study and any remaining missing data were accounted for on a pairwise basis during analyses.

Measures

Transgender Congruence Scale (TCS; Kozee et al., 2012)

The TCS is a 12-item scale containing two subscales: *Appearance Congruence* and *Gender Identity Acceptance* that measure aspects of congruence between external appearance and gender identity, as well as acceptance of gender identity, respectively. The measure was slightly adapted, with the authors' permission, to capture the individual's experiences over the prior 2-week period by changing each item to past tense. This change was made to better fit the timeframe for some of the other measures. Such adaptations have demonstrated utility

in previous applications such as a weekly assessment version of the Penn State Worry Questionnaire (PSWQ; Stoeber & Bittencourt, 1998). Higher scores indicate greater external congruence and acceptance of gender identity. The TCS has been demonstrated to be a reliable and valid measure (Kozee et al., 2012). Validity and internal consistency of the TCS for the current study are further detailed below.

Gender Minority Stress and Resilience Scale (GMSR; Testa et al., 2015)

The GMSR is a 58-item scale with 9 subscales that assess the following: (1) distal marginalization stressors: *gender-related discrimination, gender-related rejection, gender-related victimization, non-affirmation of gender identity, internalized transphobia*; (2) proximal marginalization stressors: *negative expectations for future events, non-disclosure* (of individual's gender identity or transition history); and (3) resiliency factors: *pride* (of individual's gender identity) and *TGD community connectedness*. Each of these subscales is independently scored, and higher scores on each indicate greater amount of experience or alignment with the respective domain. The GMSR has been demonstrated to be a reliable and valid measurement of marginalization stress and resilience in TGD individuals (Testa et al., 2015). In the current sample, the GMSR demonstrated acceptable to good internal consistency on most subscales, with Cronbach's α ranging from 0.65 to 0.84. However, *TGD community connectedness* demonstrated poor internal consistency ($\alpha = 0.48$); thus, it was not included in the current analysis.

Gender Preoccupation and Stability Questionnaire (GPSQ; Hakeem et al., 2016)

The GPSQ is a 14-item scale that assesses the importance, thoughts, and comfort of individual gender identity, as well as stability of one's sense of identity, and desire for medical gender affirmation. Higher scores indicate greater gender dysphoria. The GPSQ has been demonstrated to be a reliable and valid measurement of gender dysphoria (Hakeem et al., 2016). The GPSQ demonstrated acceptable internal consistency ($\alpha = 0.76$) in the current sample.

Trans Collaborations Clinical Check-In (TC³; Holt et al., 2019)

The TC³ is an 18-item scale of overall, everyday experiences of TGD individuals including confidence in handling marginalization, social support, comfort with primary and secondary physical sex characteristics, gender presentation, and disclosure of gender identity. Higher scores indicate better functioning and comfort with gender identity. The TC³ has been demonstrated to be valid and reliable (Holt et al., 2019). The TC³ demonstrated acceptable internal consistency ($\alpha = 0.77$) in the current sample.

Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001)

The PHQ-9 is a 9-item scale that assesses depressive symptomatology. Higher scores indicate greater depression. The PHQ-9 has been demonstrated to be a reliable and valid measurement of depressive symptoms (Kroenke et al., 2001) and has previously demonstrated high internal consistency in TGD samples (Bradford et al., 2019). The PHQ-9 demonstrated good internal consistency ($\alpha = 0.82$) in the current sample.

Generalized Anxiety Disorder-7 (GAD-7; Spitzer et al., 2006)

The GAD-7 is a 7-item scale that assesses generalized anxiety, but also has demonstrated sensitivity and specificity for panic disorder, social anxiety disorder, and post-traumatic stress disorder (Kroenke et al., 2007). Higher scores indicate greater anxiety. The GAD-7 has been demonstrated to be a reliable and valid measurement of anxiety symptoms (Löwe et al., 2008) and has previously demonstrated internal consistency in TGD samples (Bradford et al., 2019). The GAD-7 demonstrated acceptable internal consistency ($\alpha = 0.79$) in the current sample.

Positive and Negative Affect Schedule (PANAS; Watson et al., 1988)

The PANAS contains two subscales: *Positive Affect* and *Negative Affect* which measure dimensions of mood states. Higher scores indicate

greater positive or negative affect on their respective scales. The PANAS has been demonstrated to be a reliable and valid measurement of affective symptoms (Crawford & Henry, 2004) and various adaptations of the scale have previously been used in TGD samples (Bradford et al., 2019; McLemore, 2015). The PANAS demonstrated acceptable internal consistency for positive affect ($\alpha = 0.76$) and for negative affect ($\alpha = 0.85$) in the current sample.

Satisfaction with Life Scale (SWLS; Diener et al., 1985)

The SWLS is a 5-item measure of life satisfaction as per global judgments by the individual. Higher scores suggest greater life satisfaction. The SWLS has been demonstrated to be a reliable and valid measurement of life satisfaction (Pavot et al., 1991) and has previously demonstrated internal consistency within a TGD sample (Barr et al., 2016). The SWLS demonstrated acceptable internal consistency ($\alpha = 0.71$) in the current sample.

Data Analytic Approach

To examine the validity of the prior factor structure derived from an exploratory approach, data analyses were conducted in three phases similar to the approach outlined by Schmitt et al. (2018). First, preliminary analyses were conducted, and descriptive statistics were obtained to determine whether the data met the basic assumptions for latent variable modeling. Relying on results from a Monte Carlo estimation for minimum sample size requirements (Wolf et al., 2013), we determined we had an adequate sample size to conduct confirmatory factor analysis of the previous exploratory approach.

Second, confirmatory factor analysis was completed using the identified factor structures by Kozee et al. (2012). CFAs were completed using R (Version 3.6; R Core Team, 2018) with the *lavaan* package (Rosseel, 2012). Model fit was assessed using χ^2 , comparative fit index (CFI), root mean squared error of approximation (RMSEA), and the standardized root mean square residual (SRMR), using criteria recommended by Hu and Bentler (1998): $CFI \geq 0.95$, $RMSEA \leq 0.05$, and $SRMR \leq 0.08$.

The third phase of data analysis included re-specifying the model by identifying candidates for removal by examining poorly loading items, modification indices, and items with standardized residual covariances greater than 2 (Kenny, 2011). Once sufficient factor structure was obtained, correlational analyses were performed using SPSS, Version 22 (IBM Corp. 2013) to assess for convergence and/or divergence with other established gender identity-related and mental health-related constructs. Missing data were excluded using pairwise deletion methods to ensure that the majority of data were retained for each participant. This method was chosen as missing data methods, such as multiple imputation, have the potential to artificially skew data, due the possibility that data are not missing at random (Sterne et al., 2009). The new factor structure scores (*TCS-10 Total*, *Appearance Congruence*, and *Gender Identity Acceptance*) were compared to TGD-specific domains, including the following: GMSR pride, discrimination, rejection, victimization, non-affirmation, internalized transphobia, negative expectations for the future, and nondisclosure of gender identity; gender-related dysphoria (*GPSQ*); and gender-related functioning (*TC³*). They were also compared to general well-being measures, namely depression (*PHQ-9*), anxiety (*GAD-7*), positive and negative affect (*PANAS*), and satisfaction with life (*SWLS*). Finally, demographic variables (including gender identity categories, gender assigned at birth, and age) were assessed using ANOVA and Pearson's correlations to understand differences and/or related constructs for the final reduced *TCS-10* scores.

Results

Internal Consistency of Original Factors

Examination of Cronbach's alpha using the original factor structure of the TCS (including the *Total Score* and the two factors *Appearance Congruence Subscale* and *Gender Identity Acceptance Subscale*) indicated that the internal consistency of factors was not as strong in the current sample ($\alpha = 0.70$, $\alpha = 0.72$, $\alpha = 0.49$, respectively) compared to the original validation study ($\alpha = 0.93$, $\alpha = 0.94$, $\alpha = 0.77$, respectively) (Kozee et al., 2012). In particular, the *Gender Identity*

Acceptance subscale underperformed as a factor grouping, demonstrating poor internal consistency with the original items still intact.

Data Distribution

Normality of data distributions was examined. The skewness for individual items was between -0.65 and -0.13 , whereas the kurtosis was between -1.12 and -0.57 , which fell within the bounds of normality. As the absolute values of skewness were less than 3 and kurtosis were less than 10, the data were determined to be within acceptable ranges of normality for all three of the original TCS scales (Kline, 2005; Kozee et al., 2012).

Confirmatory Factor Analysis

Table 2 highlights model-based comparisons among the steps of CFA and item deletions.² The initial CFA examining the proposed two factor solution evidenced mixed results across model fit indices ($\chi^2(53) = 99.88$, $p < 0.001$, $\chi^2/df = 1.88$, CFI = 0.82, RMSEA = 0.07, 90% CI [0.047–0.082], and SRMR = 0.07). This finding differs from the original validation study (Kozee et al., 2012) which found their final model to demonstrate adequate fit ($\chi^2(53) = 167.41$, $p < 0.001$, $\chi^2/df = 3.16$, CFI = 0.96, RMSEA = 0.08, and SRMR = 0.04).

As the original model did not perform as well as expected, individual inter-item and item-total correlations were first examined, which revealed that several items demonstrated high correlations with other items as well as weak relationships with the total score (correlations < 0.30 ; refer to **Table 1** for item-total and inter-item correlations). Thus, a systematic approach was employed to determine specific eligibility for item reduction. All modifications were performed iteratively, and each item removed subsequently improved model fit.

Item 10 (“I was not proud of my gender identity”) from the *Gender Identity Acceptance* factor was removed because it had multiple standardized residual covariances greater than 2.0. Following removal

² In the interest of transparency, please note that in response to feedback from an anonymous reviewer on an earlier version of this manuscript, the data analyses were re-conceptualized, leading to retention of additional scale items across two factors.

of this item, a two-factor model composed of eleven items remained measuring acceptance and appearance. The resulting model evidenced improved but mixed results across model fit indices ($\chi^2(43) = 61.91$, $p < 0.031$, $\chi^2/df = 1.43$, CFI = 0.91, RMSEA = 0.05, 90% CI [0.019, 0.067], and SRMR = 0.06). After examining the standardized residual covariance matrix, item 5 (“My physical body represented my gender identity”) from the *Appearance Congruence* factor was selected for removal as it had three standardized residual covariances greater than 2.0. The resulting model evidenced excellent model fit ($\chi^2(34) = 33.05$, $p = 0.51$, $\chi^2/df = 0.97$, CFI = 1.0, RMSEA = 0.00, 90% CI [0.00, 0.005], and SRMR = 0.05). Thus, the final model was a 10-item solution, retaining the two original factors (*Appearance Congruence* and *Gender Identity Acceptance*; refer to **Tables 2 and 3**).

Convergence/Divergence

TGD-Specific Constructs

Pearson’s correlations (see **Table 4**) demonstrated that higher scores on the reduced *TCS-10 Total* scale, as well as *Appearance Congruence* and *Gender Identity Acceptance* subscales, were all associated with higher ratings of gender-related functioning (TC^3) and gender-related pride ($GMSR$). However, for all three scales, the correlations with negative expectations about the future ($GMSR$) and nondisclosure ($GMSR$) were all nonsignificant. However, there were some differences in patterns of relationships between the total and subscale scores with the remaining constructs.

Table 2 Confirmatory factor analysis model-based comparisons

Models	χ^2	χ^2/df	CFI	RMSEA	[90% CI]	SRMR	AIC	BIC
Model 1	99.88	1.88	.82	.07	[.047, 0.82]	.07	7864.12	7947.80
Model 2	61.91	1.43	.91	.05	[.019, .067]	.06	7174.54	7251.52
Model 3	33.05	.97	1.0	.00	[0.00, .005]	.05	6520.70	6591.00

Model 1, Kozee et al. original model; *Model 2*, step 1 reduced model; *Model 3*, Transgender Congruence scale, 10-item

CFI, comparative fit index; RMSEA, root mean square error of approximation; CI, confidence interval; SRMR, standardized root mean square residual; AIC, Akaike Information; BIC, Bayesian Information Criterion

Table 3 Factor loadings for original TCS and TCS-10 reduced models

Items by original factors	Item #	Original		TCS-10		Factor in TCS-10 model/ reason removed
		λ	λ	λ	λ	
Appearance Congruence (Kozee et al., 2012)						
My outward appearance represented my gender identity	1	.493	.481		Appearance Congruence	
I experienced a sense of unity between my gender identity and my body	2	.498	.447		Appearance Congruence	
My physical appearance adequately expressed my gender identity	3	.561	.604		Appearance Congruence	
I was generally comfortable with how others perceived my gender identity when they look at me	4	.624	.643		Appearance Congruence	
My physical body represented my gender identity	5	.559	–		Removed due to high covariance	
The way my body currently looks did not represent my gender identity (Reversed)	6	.276	.269		Appearance Congruence	
I was happy with the way my appearance expressed my gender identity	7	.457	.491		Appearance Congruence	
I did not feel that my appearance reflects my gender identity (Reversed)	8	.280	.316		Appearance Congruence	
I felt that my mind and body were consistent with one another	9	.436	.398		Appearance Congruence	
Gender Identity Acceptance (Kozee et al., 2012)						
I was not proud of my gender identity (Reversed)	10	.391	–		Removed due to high covariance	
I was happy that I have the gender identity that I do	11	.555	.494		Gender Identity Acceptance	
I had accepted my gender identity	12	.537	.561		Gender Identity Acceptance	

Table 4 Correlations between *TCS-10 Total* and subscales with gender and well-being constructs including means, standard deviations, and range

	<i>Total</i>	<i>AC</i>	<i>GIA</i>	<i>M</i>	<i>SD</i>	<i>Range</i>
<i>TCS-10 Total</i>	—	.96***	.50***	3.24	.60	1.70–5.00
- Appearance Congruence (<i>AC</i>)	.96***	—	.22**	3.18	.67	1.00–5.00
- Gender Identity Acceptance (<i>GIA</i>)	.50***	.22**	—	2.34	.61	1.00–3.33
<i>TC</i> ³	.58***	.54***	.33***	56.28	9.18	28.00–89.00
<i>GPSQ</i>	-.23**	-.18*	-.26***	41.83	8.20	15.00–62.00
<i>GMSR</i>						
- Discrimination	-.19**	-.11	-.29***	3.63	1.69	0.00–5.00
- Rejection	-.12	-.06	-.21**	4.15	2.06	0.00–6.00
- Victimization	-.04	.05	-.26***	3.99	2.26	0.00–6.00
- Non-affirmation	-.25***	-.28***	.00	14.14	4.69	0.00–24.00
- Internalized transphobia	-.18**	-.09	-.34***	16.77	6.75	0.00–30.00
- Pride	.36***	.26***	.44***	18.66	5.84	4.00–32.00
- Negative expectations	-.02	-.03	.02	20.25	6.52	0.00–36.00
- Nondisclosure of gender identity	.09	.13	-.07	10.98	4.06	0.00–20.00
<i>PHQ-9</i>	-.11	-.05	-.22**	13.01	5.81	0.00–27.00
<i>GAD-7</i>	-.02	.02	-.11	10.74	4.60	0.00–21.00
<i>PANAS</i>						
- Positive	.31***	.26***	.25***	28.90	6.23	11.00–45.00
- Negative	-.03	.03	-.19**	26.17	7.42	9.00–40.00
<i>SWLS</i>	.43***	.43***	.17*	21.59	5.77	6.00–35.00

TCS, Transgender Congruence Scale; *TC*³, Trans Collaborations Clinical Check-in; *GPSQ*, Gender Preoccupation and Stability Questionnaire; *GMSR*, Gender Minority Stress and Resilience; *PHQ-9*, Patient Health Questionnaire (9-item); *GAD-7*, Generalized Anxiety Disorder (7-item); *PANAS*, Positive and Negative Affect Scale; *SWLS*, Satisfaction with Life Scale.

N's range from 202 to 208 due to missing data

*** $p < .001$; ** $p < .01$; * $p < .05$

First, higher *TCS-10 Total* scores were associated with lower ratings of gender-related dysphoria (*GPSQ*), non-affirmation of gender identity (*GMSR*), internalized transphobia (*GMSR*), and gender-related discrimination (*GMSR*). However, the correlations between *TCS-10 Total* scores and rejection (*GMSR*) and victimization (*GMSR*) were not significant. Higher *TCS-10 Appearance Congruence* scores were significantly associated with lower ratings of gender-related dysphoria and nonaffirmation; however, there were no significant relationships with discrimination, rejection, victimization, nor internalized transphobia. There was also no significant relationship for *TCS-10 Gender Identity Acceptance* with non-affirmation of gender identity.

General Well-being Constructs

For measures of general well-being, higher scores on *TCS-10 Total* scale, *Appearance Congruence*, and *Gender Identity Acceptance* revised scores were related to greater positive affect (*PANAS*) and satisfaction with life (*SWLS*). However, for all three *TCS-10* scales, there was no significant relationship with anxiety (*GAD-7*). Again, there were differences in relationships for the remaining constructs with the three revised scale scores. For both the *TCS-10 Total* and *Appearance Congruence* scores, correlations with ratings of depression (*PHQ-9*) and negative affect (*PANAS*) were nonsignificant. Finally, higher *TCS-10 Gender Identity Acceptance* scores were significantly related to lower depression and negative affect.

Demographic Comparisons

For those who specified within one of the 3-category gender constructs (*transmasculine*, *transfeminine*, *gender diverse*), ANOVA comparisons of the *TCS-10 Total* score revealed a significant effect of gender identity category ($F(2,192) = 4.07$, $Mse = 0.350$, $p = 0.019$). This excluded the 15 participants who did not specify within one of these categories. Follow-up comparisons using Tukey's LSD demonstrated significantly lower *TCS-10 Total* scores for those who identify as gender diverse ($M = 3.017$, $SD = 0.584$) compared to both transmasculine ($M = 3.291$, $SD = 0.599$, $p = 0.015$) and transfeminine ($M = 3.312$, $SD = 0.589$, $p = 0.009$) identified groups with no difference between the latter two groups ($p = 0.825$). There was a similar pattern of relationships for the *TCS-10 Appearance Congruence* subscale ($F(2,192) = 6.00$, $Mse = 0.422$, $p = 0.003$), with significantly lower scores for gender diverse identities ($M = 2.875$, $SD = 0.683$) compared to both transmasculine ($M = 3.233$, $SD = 0.647$, $p = 0.004$) and transfeminine ($M = 3.272$, $SD = 0.631$, $p = 0.001$) identities. This is contrary to prior findings that there were no significant differences in *TCS Appearance Congruence* scores between gender diverse and transmasculine and transfeminine identified groups (Jones et al., 2019b). There were no significant differences between groups for *TCS-10 Gender Identity Acceptance* scores ($F(2,192) = 0.213$, $Mse = 0.385$, $p = 0.808$). Age was not correlated with *TCS-10 Total* scores ($r(156) = -0.082$, p

= 0.309), but as noted above, there was substantial missing data for age. ANOVA analyses revealed no significant mean difference in *TCS-10 Total* scores between gender assigned at birth (i.e., male vs. female; $F(1,193) = 0.127$, $Mse = 0.364$, $p = 0.722$), nor for the *TCS-10 Appearance Congruence* ($F(1,193) = 0.350$, $Mse = 0.451$, $p = 0.555$) or *TCS-10 Gender Identity Acceptance* ($F(1,193) = 0.300$, $Mse = 0.389$, $p = 0.585$) subscales.

Discussion

The current study examined the two-factor structure of the TCS found in the original validation study by Kozee et al. (2012) in an additional sample of TGD individuals with a larger proportion of participants who identified outside the gender binary. Overall, the results indicated that the fit of the original model did not perform as well in the current sample compared to the initial studies, and that internal consistency was lower among the three scale groupings, which was contrary to the expectation that the original factors and items would be retained in the current sample and would demonstrate internal consistency. While we retained the original two-factor structure (*Appearance Congruence* and *Gender Identity Acceptance*), removal of two items via a series of stepwise analyses greatly improved model fit in the current sample, resulting in a 10-item reduced model (*TCS-10*).

As hypothesized, *TCS-10 Total*, *Appearance Congruence*, and *Gender Identity Acceptance* scores also demonstrated convergence in the expected direction with several TGD-specific measures. There were positive correlations with gender-related well-being and pride, as well as negative correlations with gender dysphoria, internalized transphobia, non-affirmation of gender identity, and gender-related discrimination, rejection, and victimization. However, there were also inconsistencies in relationships between constructs (e.g., *TCS-10* and aforementioned constructs) that demonstrated different patterns of results than expected, warranting further consideration. For example, only *TCS-10 Gender Identity Acceptance* was related to rejection and victimization, two distal marginalization stress factors in the Testa et al. (2015) model, while the remaining two distal factors of gender-related discrimination and non-affirmation of gender identity were

related to the *TCS-10 Total* and *Gender Identity Acceptance* and the *TCS-10 Total* and *Appearance Congruence* scales, respectively. It stands to reason experiences of discrimination, rejection, and victimization relate to one's internal acceptance of gender (Rood et al., 2017). These findings also suggest that there is a greater likelihood that those with lower appearance congruence might be more likely have to constantly reiterate to others or behave in ways to portray and affirm their gender identity. This highlights the need to support and affirm TGD people who are less comfortable or feel less congruent in their gender as there is a greater likelihood for negative experiences and discrimination to relate to poorer self-image. For some individuals, this support may include medical/surgical affirmation procedures, but all TGD individuals would benefit from dismantling oppressive systemic structures that devalue diverse gender identities (Cizek et al., 2021).

The lack of relationship of the total score with rejection and victimization was surprising but may have been due to the fact that *Appearance Congruence* subscale contributes a majority of items to the total score and was itself not related to rejection and victimization. Non-affirmation items in the GMSR (e.g., "I have difficulty being perceived as my gender") pertain more to physical perceptions of gender, rather than internal views of gender identity, making the pattern of relationships with the *TCS-10 Total* and *Appearance Congruence* scales consistent with content validity of the items in the TCS. Furthermore, contrary to our hypothesis, two of the three proximal marginalization stressors (negative expectations and nondisclosure) in the Testa et al. model were unrelated to the *TCS-10 Total* and both subscales, with the third, internalized transphobia, having a modest relationship with the total score and a stronger relationship with *Gender Identity Acceptance*. This suggests that one's sense of acceptance of their own internal experience of gender may be more strongly related to internalized beliefs about the meaning of being TGD rather than to external gender congruence. Additionally, both internal and external experiences of gender may be somewhat removed from what it means for how the world will react to one's gender identity. Alternatively, it may be that this scale now measures more about appearance rather than internal gender concepts given the removal of one of the three items that assessed personal acceptance of participants' gender identities.

Consistent with hypothesized relationships between the TCS and general well-being measures, higher *TCS-10 Total* scores were associated with lower gender dysphoria, a TGD-specific well-being measure, and greater overall life satisfaction and positive affect. Contrary to previous findings by Kozee et al. (2012), *TCS-10 Total* and subscale scores did not correlate with anxiety; however, *Gender Identity Acceptance* was negatively related to depression and negative affect. This differs from the original TCS validation studies in which Kozee et al. (2012) reported correlations between all three TCS scales (*Total*, *Appearance Congruence*, *Gender Identity Acceptance*) with both depression ($r = -0.53, -0.49, -0.41$, respectively) and anxiety ($r = -0.33, -0.30, -0.24$, respectively) using the Beck Depression Inventory-II (BDI-II; Beck et al., 1996) and the Beck Anxiety Inventory (BAI; Beck & Steer, 1993), respectively. It is unlikely that the different pattern of results is attributable to the specific scales as the PHQ-9 and BDI-II tend to correlate highly (e.g., $r = 0.77$; Kung et al., 2013). Similarly, the GAD-7 and the BAI correlate with one another (e.g., $r = 0.72$; Spitzer et al., 2006).

The lack of correlation between *TCS-10* scales and wellbeing measures may have been impacted by the removal of items from the original item structure for the *TCS Total* scale and *Appearance Congruence* subscale. In order to understand the changes that item removal may have made for the differences in patterns of relationships between the original TCS in the validation study and the *TCS-10* in the current study, the authors conducted additional post hoc analyses to explore any changes in relationships before and after item removal. Specifically, there were modest but significant relationships between the original TCS total score and negative affect ($r = 0.17, p = 0.017$), as well as for the original Appearance Congruence subscale with anxiety ($r = 0.15, p = 0.033$) and negative affect ($r = 0.18, p = 0.011$) in the current sample prior to item removal. This suggests that the removed items may impact relationships between the TCS and negative affect as well as anxiety; however, it should be noted that these effects were small prior to item removal. However, this did not impact relationships with the original *Gender Identity Acceptance* items as they were not significantly related to constructs of well-being in the present sample prior to item removal ($r = 0.02-0.05, p > 0.05$). Due to some clear differences in relationships among these variables, future

research is needed to better understand the relationship between gender congruence and negative affective experiences.

Finally, contrary to hypothesis that scores would not be significantly different between transmasculine, transfeminine, and gender diverse identity groups, there was a significant difference in average scores between the binary and gender diverse gender groups. Those with gender diverse identities scored significantly lower on the *Total* and *Appearance Congruence* scales than those with transmasculine or transfeminine gender identities. While this is not consistent with expectations for the TCS to perform equally well for those across the gender spectrum, it may be that those with gender diverse identities are less likely to endorse appearance-related congruence as they have difficulty expressing a nonbinary identity that lacks clear social norms or, perhaps, they do not aspire to “fit” within a binary gender category. Both of these explanations would be likely to result in lower appearance-related ratings. However, as hypothesized, the groups scored equivalently on the *Gender Identity Acceptance* highlighting that overall, the sample scored similarly on items pertaining to internal acceptance of their own gender identity. As such, the overall TCS appears to capture internal acceptance well across the gender spectrum; however, there may be further investigation needed in order to understand how gender diverse individuals read and understand aspects of external gender congruence within the TCS.

Overall, this study demonstrated that in general, gender congruence is associated with greater TGD resilience and well-being, especially pertaining to internal views of gender. This is consistent with theoretical models that emphasize resilience, including positive affect, is at least partially orthogonal to stress, negative affect, or mental health rather than being an opposite pole (e.g., Breslow et al., 2015). High resilience is not simply low stress or better mental health. It also highlighted that experiences of marginalization may well have implications for internal views of gender as well as associations with negative mental health. As such, gender congruence appears to be an important aspect for TGD resiliency and should be considered in future research on TGD health disparities. While the ideal is for communities and institutions to address disparities and discrimination at the source, in the meantime, a salient need exists for building resiliency in the face of marginalizing experiences and

to build on gender diverse identity affirmation. Further, as there are disparate findings regarding differences in gender congruence among gender binary and gender diverse people, there is a need for further understanding regarding gender identity and congruence among gender diverse communities.

Limitations

The results of this study must be considered in light of its limitations. First, the primary confirmatory factor analysis was collapsed across gender identity groups. Individuals who identified as gender diverse made up about one-fifth of the sample and scored significantly lower on *TCS-10 Total* and *Appearance Congruence* constructs than those who identified as transfeminine or transmasculine. Participants in the gender diverse group used a variety of labels for themselves in the free response option, suggesting heterogeneity in gender identity. Gender congruence may have many meanings for this group which may not be measured equally well as when the same scale as is used for TGD people who identify as either transmasculine or transfeminine, particularly in appearance-related aspects. Since gender groups were not compared in the original validation samples (Kozee et al., 2012) and there were disparate findings for those with gender diverse identities, further study is needed to test the reliability and validity of the TCS or *TCS-10* in a primarily gender diverse sample.

Secondly, while the sample was sufficiently large to detect appropriate model solutions for similar CFAs (Wolf et al., 2013), replications are needed to confirm the current factor analysis. Relatedly, as the data were collected from one sample, additional comparison samples would allow for examination of the replicability of the current factor analysis. A third limitation is that the sample did not represent the diversity of TGD communities which limits generalizability. The sample was mostly White or European American (75%), which highlights the need for further examination among other salient identity factors such as race and ethnicity, as well as other facets of identities. A larger, more diverse sample would allow for greater statistical power to perform comparisons among demographic and identity factors as well as more complex structural analyses. Our research team is currently investigating research

methodologies and recruitment strategies that yield samples that are more representative of the diversity of TGD communities, particularly communities of color.

Additionally, as with most general recruitment for online survey participation, there is a self-selection bias that must be considered. Due to this type of recruitment, the findings likely are not entirely representative of all TGD people and experiences and should be interpreted in the context of the current sample who chose to participate in the study. This self-selection process limits the ability to understand the rate of response and potential communities that might be underrepresented in the sample as mentioned previously. Future research might find ways to track reach and rate of response to understand this process more thoroughly of who does and does not choose to participate in survey research.

A final limitation was the adaptation of the items to past tense (i.e., *My outward appearance represented my gender identity*) deviates from the original items that were phrased in current tense (i.e., *My outward appearance represents my gender identity*). This change was made to better fit with other measures in the overall survey data collection for the sake of consistency of timeframe. Given that the original instructions also included the 2-week reference, it seems unlikely that this grammatical change impacted the results. However, such a possibility cannot be completely ruled out.

Conclusions

Measurement of key constructs is crucial to the scientific endeavor. Empirical research to improve quality of life and reduce health disparities for TGD people cannot proceed without culturally appropriate and sound psychometric measures. Kozee et al.'s (2012) work to develop the TCS is important because it addressed the need for empirically supported measurement and avoided a narrow focus on transition and gender binary assumptions for TGD identities. The current study failed to support the adequate model fit of the 12-item TCS, but the current findings suggest that a 10-item, two-factor measure is an important and valid measure of gender congruence that addresses various aspects of TGD-identity congruence domains.

Implications for Social Policy and Research

It is crucial to contextualize well-being of TGD persons within a socio-ecological framework, including experiences of marginalization, as they contribute to overall health disparities among gender minority individuals (Williams & Mann, 2017). The ultimate onus of change regarding negative perceptions of gender diverse identities needs to take place at systems and policy levels, including protections for the rights of all identities and inclusion of all persons, namely gender diverse individuals. While the current analysis focuses on individual levels of functioning and resilience, ultimate change must happen at these broader levels in order to alleviate the burdens faced by marginalized communities.

Gender congruence should be conceptualized as a key aspect of TGD experiences in future research given its role in the overall well-being of TGD people. However, further attention is needed for replication and convergence of findings in diverse samples to understand the nuances and complexities of intersectional identities that impact experiences of gender identity (Watson et al., 2020) that were beyond the bounds of the current sample. Overall, the importance of gender congruence as an aspect of resilience spotlights the strength of TGD persons when much of the scientific and cultural discourse is about deficits.

* * * *

Acknowledgments We would also like to acknowledge the contributions of Jolene Smyth, the Trans Collaborations Local Community Board, and the participants for taking the time to contribute to our work.

Author Contributions T. Zachary Huit was the primary author, with significant contributions regarding content from Allura L. Ralston and Debra A. Hope and analytical approach from J. Kyle Haws. Natalie R. Holt was responsible for the recruitment of participants and collection of data via the logistical support of Jae Puckett. Additionally, all authors apart from the third author contributed significantly to content expertise regarding TGD communities and the approach of the manuscript, including Richard A. Mocarski. Finally, Nathan Woodruff served as the community leadership voice for the Trans Collaborations Local Community Board to approve content and study approach.

Funding This work was supported, in part, by grants from the National Institute of Mental Health (R21MH108897), Great Plains IDEa-CTR (U54GM115458) pilot award, and University of Nebraska Systems Science Team Building Award.

Availability of Data and Material Due to the sensitive nature of questions and identities contained in the data set with a vulnerable TGD populations, the research team and participants, under the advisement of community board members, have agreed to not make case-level data publicly available. Participants are ensured that only aggregate data will be presented in research dissemination to ensure confidentiality and anonymity.

Code Availability The code is available upon request from the authors.

Conflict of Interest The authors declare there are no competing interests.

References

- Beck, A. T., & Steer, R. A. (1993). *Beck Anxiety Inventory Manual*. The Psychological Corporation Harcourt Brace.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. Psychological Corporation.
- Benson, K. E. (2013). Seeking support: Transgender client experiences with mental health services. *Journal of Feminist Family Therapy*, 25(1), 17–40. <https://doi.org/10.1080/08952833.2013.755081>
- Barr, S. M., Budge, S. L., & Adelson, J. L. (2016). Transgender community belongingness as a mediator between strength of transgender identity and well-being. *Journal of Counseling Psychology*, 63(1), 87–97. <https://doi.org/10.1037/cou0000127>
- Bradford, N. J., Rider, G. N., & Spencer, G. N. (2019). Hair removal and psychological well-being in transfeminine adults: Associations with gender dysphoria and gender euphoria. *Journal of Dermatological Treatment*, 1–8 <https://doi.org/10.1080/09546634.2019.1687823>
- 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>
- Chang, T. K., & Chung, Y. B. (2015). Transgender microaggressions: Complexity of the heterogeneity of transgender identities. *Journal of LGBT Issues in Counseling*, 9(3), 217–234. <https://doi.org/10.1080/15538605.2015.1068146>
- Chodzen, G., Hidalgo, M. A., Chen, D., & Garofalo, R. (2019). Minority stress factors associated with depression and anxiety among transgender and gender-nonconforming youth. *Journal of Adolescent Health*, 64(4), 467–471. <https://doi.org/10.1016/j.jadohealth.2018.07.006>

- Ciszek, E., Mocarski, R., Price, S., & Almeida, E. (2021). Discursive stickiness: Affective institutional texts and activist resistance. *Public Relations Inquiry*, 1–16. <https://doi.org/10.1177/2F2046147X211008388>
- Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43(3), 245–265. <https://doi.org/10.1348/0144665031752934>
- Diener, E., Emmons, R., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49, 71–75. <https://doi.org/10.1207/s15327752jpa490113>
- de Vries, K. M. (2012). Intersectional identities and conceptions of the self: The experience of transgender people. *Symbolic Interaction*, 35(1), 49–67. <https://doi.org/10.1002/symb.2>
- Hakeem, A., Črnčec, R., Asghari-Fard, M., Harte, F., & Eapen, V. (2016). Development and validation of a measure for assessing gender dysphoria in adults: The Gender Preoccupation and Stability Questionnaire. *International Journal of Transgenderism*, 17(3–4), 131–140. <https://doi.org/10.1080/15532739.2016.1217812>
- Holt, N. R., Huit, T. Z., Shulman, G. P., Meza, J. L., Smyth, J. D., Woodruff, N., Mocarski, R., Puckett, J. A., & Hope, D. A. (2019). Trans Collaborations Clinical Check-In (TC³): Initial validation of a clinical measure for transgender and gender diverse adults receiving psychological services. *Behavior Therapy*, 50(6), 1136–1149. <https://doi.org/10.1016/j.beth.2019.04.001>
- Hope, D. A., Mocarski, R., Bautista, C. L., & Holt, N. R. (2016). Culturally competent evidence-based behavioral health services for the transgender community: Progress and challenges. *American Journal of Orthopsychiatry*, 86(4), 361–365. <https://doi.org/10.1037/orto000197>
- Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3(4), 424–453. <https://doi.org/10.1037/1082-989X.3.4.424>
- IBM Corp. (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.
- Jackman, K. B., Dolezal, C., Levin, B., Honig, J. C., & Bockting, W. O. (2018). Stigma, gender dysphoria, and nonsuicidal self-injury in a community sample of transgender individuals. *Psychiatry Research*, 269, 602–609. <https://doi.org/10.1016/j.psychres.2018.08.092>
- Jones, B. A., Bouman, W. P., Haycraft, E., & Arcelus, J. (2019a). The Gender Congruence and Life Satisfaction Scale (GCLS): Development and validation of a scale to measure outcomes from transgender health services. *International Journal of Transgenderism*, 20(1), 63–80. <https://doi.org/10.1080/15532739.2018.1453425>
- Jones, B. A., Bouman, W. P., Haycraft, E., & Arcelus, J. (2019b). Gender congruence and body satisfaction in nonbinary transgender people: A case control study. *International Journal of Transgenderism*, 20(2–3), 263–274. <https://doi.org/10.1080/15532739.2018.1538840>

- Kenny, D. A. (2011). *Respecification of latent variable models*. Retrieved 12 Jan 2020, from <https://davidakenny.net/cm/respec.htm>
- King, M., Dinos, S., Shaw, J., Watson, R., Stevens, S., Passeti, F., Weich, S., & Serfaty, M. (2007). The Stigma Scale: Development of a standardised measure of the stigma of mental illness. *The British Journal of Psychiatry*, *190*, 248–254. <https://doi.org/10.1192/bjp.bp.106.024638>
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford.
- Kozee, H. B., Tylka, T. L., & Bauerband, L. A. (2012). Measuring transgender individuals' comfort with gender identity and appearance: Development and validation of the Transgender Congruence Scale. *Psychology of Women Quarterly*, *36*(2), 179–196. <https://doi.org/10.1177/0361684312442161>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine*, *16*(9), 606–613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Kroenke, K., Spitzer, R. L., Williams, J. B., Monahan, P. O., & Löwe, B. (2007). Anxiety disorders in primary care: Prevalence, impairment, comorbidity, and detection. *Annals of Internal Medicine*, *146*(5), 317–325. <https://doi.org/10.7326/0003-4819-146-5-200703060-00004>
- Kung, S., Alarcon, R. D., Williams, M. D., Poppe, K. A., Moore, M. J., & Frye, M. A. (2013). Comparing the Beck Depression Inventory-II (BDI-II) and Patient Health Questionnaire (PHQ-9) depression measures in an integrated mood disorders practice. *Journal of Affective Disorders*, *145*(3), 341–343. <https://doi.org/10.1016/j.jad.2012.08.017>
- Kuper, L. E., Nussbaum, R., & Mustanski, B. (2012). Exploring the diversity of gender and sexual orientation identities in an online sample of transgender individuals. *Journal of Sex Research*, *49*(2–3), 244–254. <https://doi.org/10.1080/00224499.2011.596954>
- Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., & Herzberg, P. Y. (2008). Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Medical Care*, *46*(3), 266–274. <https://doi.org/10.1097/MLR.0b013e318160d093>
- McLemore, K. A. (2015). Experiences with misgendering: Identity misclassification of transgender spectrum individuals. *Self and Identity*, *14*(1), 51–74. <https://doi.org/10.1080/15298868.2014.950691>
- Mizock, L., & Mueser, K. T. (2014). Employment, mental health, internalized stigma, and coping with transphobia among transgender individuals. *Psychology of Sexual Orientation and Gender Diversity*, *1*(2), 146–158. <https://doi.org/10.1037/sgd0000029>
- Obasi, S. N. (2016). Naming patterns in rural south-central. *Nebraska Names*, *64*(3), 158–165. <https://doi.org/10.1080/00277738.2016.1197644>
- Pavot, W., Diener, E. D., Colvin, C. R., & Sandvik, E. (1991). Further validation of the Satisfaction with Life Scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment*, *57*(1), 149–161. https://doi.org/10.1207/s15327752jpa5701_17

- Puckett, J. A., Cleary, P., Rossman, K., Mustanski, B., & Newcomb, M. E. (2018). Barriers to gender-affirming care for transgender and gender nonconforming individuals. *Sexuality Research and Social Policy*, 15(1), 48–59. <https://doi.org/10.1007/s13178-017-0295-8>
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rood, B. A., Reisner, S. L., Puckett, J. A., Surace, F. I., Berman, A. K., & Pantalone, D. W. (2017). Internalized transphobia: Exploring perceptions of social messages in transgender and gender-nonconforming adults. *International Journal of Transgenderism*, 18(4), 411–426. <https://doi.org/10.1080/15532739.2017.1329048>
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- Schmitt, T. A., Sass, D. A., Chappelle, W., & Thompson, W. (2018). Selecting the “best” factor structure and moving measurement validation forward: An illustration. *Journal of Personality Assessment*, 100(4), 345–362. <https://doi.org/10.1080/00223891.2018.1449116>
- Shulman, G. P., Holt, N. R., Hope, D. A., MocarSKI, R., Eyer, J., & Woodruff, N. (2017). A review of contemporary assessment tools for use with transgender and gender nonconforming adults. *Psychology of Sexual Orientation and Gender Diversity*, 4(3), 304–313. <https://doi.org/10.1037/2Fsgd0000233>
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, 166(10), 1092–1097. <https://doi.org/10.1001/archinte.166.10.1092>
- Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls *BMJ* 2009;338:b2393. <https://doi.org/10.1136/bmj.b2393>
- Stoeber, J., & Bittencourt, J. (1998). Weekly assessment of worry: An adaptation of the Penn State Worry Questionnaire for monitoring changes during treatment. *Behaviour Research and Therapy*, 36(6), 645–656. [https://doi.org/10.1016/S0005-7967\(98\)00031-X](https://doi.org/10.1016/S0005-7967(98)00031-X)
- Testa, R. J., Habarth, J., Peta, J., Balsam, K., & Bockting, W. (2015). Development of the Gender Minority Stress and Resilience Measure. *Psychology of Sexual Orientation and Gender Diversity*, 2(1), 65–77. <https://doi.org/10.1037/sgd0000081>
- van den Brink, F., Vollmann, M., & van Weelie, S. (2019). Relationships between transgender congruence, gender identity rumination, and self-esteem in transgender and gender-nonconforming individuals. *Psychology of Sexual Orientation and Gender Diversity*. <https://doi.org/10.1037/sgd0000357>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>

- Watson, R. J., Wheldon, C. W., & Puhl, R. M. (2020). Evidence of diverse identities in a large national sample of sexual and gender minority adolescents. *Journal of Research on Adolescence*, 30, 431-442. <https://doi.org/10.1111/jora.12488>
- Williams, S. L., & Mann, A. K. (2017). Sexual and gender minority health disparities as a social issue: How stigma and intergroup relations can explain and reduce health disparities. *Journal of Social Issues*, 73(3), 450-461. <https://doi.org/10.1111/josi.12225>
- Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement*, 76, 913-934. <https://doi.org/10.1177/0013164413495237>