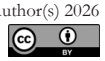
Check for updates

Group Processes & Intergroup Relations

Article

Group Processes & Intergroup Relations
1–25
© The Author(s) 2026

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13684302261499271
journals.sagepub.com/home/gpi


The Brief Transprejudice Scale: Development and Validation

Diana Cheso[1], R. Thora Bjornsdottir[1,2], Ryan McKay[1] and Hanna Zagefka[1]

Abstract

In recent years, research on predictors of transprejudice has increased, yet existing psychometric measures of the construct have limitations, including outdated terminology, conflation of gender atypicality with trans identity and a lack of multidimensional structure. We report on the development and validation of the Brief Transprejudice Scale (BTS), a concise yet multifaceted measure of transprejudice. In Study 1 (N = 341), exploratory factor analysis and Rasch analysis were used to refine an initial item pool, resulting in a three-factor (negative affect, category exclusion and moralism) 14-item scale . In Study 2 (N = 391), confirmatory factor analysis supported this factor structure, and validity tests demonstrated the scale's strong concurrent, convergent and divergent validity. Despite its brevity, this new scale provides a psychometrically sound and efficient tool that allows researchers to capture multiple aspects of transprejudice.

Keywords: transprejudice, transphobia, scale development, scale validation

Received: March 28, 2025; accepted: October 6, 2026

Introduction

Transgender individuals face persistent discrimination and hostility, with reports of trans-related hate crimes rising in recent years (Amnesty International, 2023; United Nations, 2023). In the UK, for example, recorded transphobic hate crimes increased more than sixfold between 2012 and 2022 (Metropolitan Police, 2022), and transgender people continue to face disproportionate obstacles accessing legal services and medical care (Bauer et al., 2014; Davidson et al., 2026). Underlying these behaviours and systemic barriers are transprejudice attitudes – negative beliefs, emotions and moral judgments toward transgender people. Understanding and measuring these

[1]Department of Psychology, Royal Holloway, University of London, UK
[2]Division of Psychology, University of Stirling, UK

Corresponding Author:
Diana Cheso, Department of Psychology, Royal Holloway, University of London, Egham, TW20 0EX, UK.
Email: diana.cheso@rhul.ac.uk

attitudes is therefore crucial, both for tracking changes in public attitudes and for identifying psychological predictors of bias against transgender people.

This rise in transprejudice (also called transphobia)[1] has coincided with an increase in trans-related research, which includes the development of psychometric tools to measure transprejudice (for a systematic review, see Morrison et al., 2017). Over the last two decades, a few transprejudice measures have been validated, making valuable contributions to the research on transprejudice (see Table 1 for an overview of transprejudice scales). These scales have been instrumental in identifying various predictors of transprejudice, including heightened religiosity (Acker, 2017; Elischberger et al., 2018), social dominance orientation (SDO), right-wing authoritarianism (RWA; Hatch et al., 2022; Perez-Arche & Miller, 2021) and gender/sex essentialism (Ching et al., 2020; Morgenroth et al., 2021).

Before discussing these multi-item measures, it is important to first acknowledge the use of single-item scales. Given that researchers often contend with finance- or time-related costs, single-item scales are often used to capture prejudice, transprejudice included. Popular among these are, for instance, feeling thermometers, where participants are asked to indicate their level of warmth towards transgender people (e.g., Glazier et al., 2021). Single-item measures such as these have been shown to reliably capture prejudice, albeit limited to a single dimension (e.g., affective in the case of feeling thermometers; Allen et al., 2022). However, in cases where it is necessary to explore more nuanced or complex aspects of prejudice, multi-item scales may be better suited.

As highlighted, numerous multi-item scales have been developed to capture transprejudice. It is, however, vital to continually update psychometric tools in line with societal shifts to ensure that the constructs they measure are captured accurately. Evidence suggests that there is a need for such an update. First of all, some of the previously validated scales refer to trans people using terminology that is now outdated and/or inaccurate. The most up-to-date definition of a transgender woman is someone whose sex assigned at birth was male but whose gender identity is woman, and a transgender man is someone whose sex assigned at birth was female but whose gender identity is man (Stonewall, 2024). Both the Genderism and Transphobia Scale (GTS; Hill & Willoughby, 2005) and the Attitudes Toward Transgendered Individuals Scale (Walch et al., 2012) use terms such as *transgenderists* and *transgendered*, which are now widely considered derogatory (Glaad, 2024). Thus, it is necessary to develop psychometric measures with updated language. Additionally, some of the items may inadvertently cause feelings of distress (e.g., 'I have beaten up men who act like sissies'; Hill & Willoughby, 2005). It is recommended that scale items of psychometric measures avoid, where possible, being distressing or insensitive (Allen et al., 2022).

Moreover, some of the validated scales conflate attitudes towards transgender identity with attitudes not necessarily related to this identity. For example, the GTS was developed to measure 'violence, harassment, and discrimination toward cross-dressers, transgenderists, and transexuals' (Hill & Willoughby, 2005, p. 531). In addition to the use of now outdated language to describe gender-diverse people (discussed below), it is important to highlight that individuals who wear clothing atypical of their gender are not necessarily transgender (Billard, 2018a; Glaad, 2024). Additionally, none of the 32 items in the GTS refer to transgender individuals or identity specifically; instead, items refer to 'feminine men/boys' and 'masculine women/girls' (e.g., 'Feminine men make me feel uncomfortable', 'I can't understand why a woman would act masculine'). This focus on gender atypicality rather than transgender identity may detrimentally impact the accuracy of this scale in assessing attitudes towards trans women and men.

In a similar vein, the nine-item Transphobia Scale (Nagoshi et al., 2008) has items such as 'I think there is something wrong with a person who says they are neither a man nor a woman', 'I believe that the male/female dichotomy is

Table 1

Published Scales Measuring Transprejudice

Author(s)	Measure	Overview/limitations
1. Hill and Willoughby (2005)	Genderism and Transphobia Scale (GTS)	A 32-item scale that captures attitudes towards gender atypicality but not necessarily specific to trans identity. Some items may now be considered insensitive.
2. Tee and Hegarty (2006)	Support for Trans Persons' Civil Rights Scale	Scale items focus specifically on medicalised aspects of trans identity, that is, trans people who seek gender-affirming care.
3. Nagoshi et al. (2008)	The Transphobia Scale (TS)	Scale items focus primarily on the physical aspects of trans identity, that is, gender typicality.
4. King et al. (2009)	Chinese Attitudes Towards Transgenderism and Transgender Civil Rights Scale (CATTCRS)	Many of the scale items are worded specifically for a Chinese sample.
5. Walch et al. (2012)	The Attitudes Toward Transgendered Individuals Scale (ATTI)	Scale items refer to trans people using terminology that is now outdated, that is, transgendered/transgenderism.
6. Case and Stewart (2013)	Transexual Prejudice Scale	Scale items focus specifically on trans people seeking gender-affirming care.
7. Gazzola and Morrison (2014)	Stereotypes of transgender men and transgender women	A focus on participants' endorsed stereotype content (both positive and negative) and their perceptions of traits of trans women and men.
8. Tebbe et al. (2014)	Genderism and Transphobia Scale revised and short-form versions (GTS-R; GTS-SF)	Revised and short-form versions of the previously validated GTS (Hill & Willoughby, 2005).
9. Jami and Kamal (2015)	Attitudes Toward Hijra Scale	Measures attitudes toward transgender people in a context that is culturally specific to South Asia, in this case, Pakistan.
10. Kanamori et al. (2017)	Transgender Attitudes and Beliefs Scale (TABS)	A 29-item scale. Authors purposefully oversampled Evangelical Christians in both the scale development and validation stages. Over 50% of participants reported Christian affiliation, most likely reducing the generalisability of this scale.
11. Billard (2018a)	Attitudes Toward Transgender Men and Women Scale (ATTMW)	A 24-item unidimensional scale.
12. Clark and Hughto (2020)	Transgender Knowledge, Attitudes, and Beliefs Scale (T-KAB)	A 22-item scale measuring knowledge in addition to prejudice.
13. Amigo-Ventureira et al. (2025)	Homophobia and Transphobia Test	A scale developed and validated to measure both homophobia and transprejudice.
14. Jones et al. (2024)	Trans Attitudes and Beliefs Scale (TABS) – updated	A 29-item measure. Authors updated the language of the original TABS scale (Kanamori et al., 2017).

natural'. Whilst some of the items may indeed tap into attitudes towards trans women and men, this measure may more accurately capture attitudes specifically toward nonbinary individuals (who fall under the broader trans umbrella). Indeed, the authors state that their scale was adapted from questions that 'assess a person's degree of discomfort when encountering individuals who don't conform to conventional gender norms' (Nagoshi et al., 2008, p. 523). Although capturing these aspects of trans identity attitudes is important, it is also necessary to capture trans-related attitudes outside gender expression or gender typicality.

Two existing scales validated in the last 5–10 years have incorporated more up-to-date language and appear to have implemented many of the best practice recommendations in scale development and validation (Morrison et al., 2017). These are the Attitudes Toward Transgender Men and Women Scale (ATTMW; Billard, 2018a) and the Transgender Knowledge, Attitudes and Beliefs Scale (T-KAB; Clark & Hughto, 2020). When developing the ATTMW scale, Billard (2018a) conducted qualitative interviews with general members of the public as part of the item generation stage. In a similar vein, Clark and Hughto (2020) sought input from trans healthcare experts, two of whom were trans themselves. However, in addition to being unidimensional (at least in the case of ATTMW), the resulting scales are quite long (24 items in the ATTMW and 22 items in the T-KAB); short scales are often more versatile. Additionally, the T-KAB is not purely a prejudice measure but captures trans-related knowledge as well. We therefore sought to build on this work by creating a brief yet multifactor scale with a focus on capturing prejudice.

The Present Research

Our main aim was to create an updated and brief multi-item measure of transprejudice. Indeed, measures of intergroup prejudice must routinely be updated to capture changes in the cultural zeitgeist (Billard, 2018b; Jones et al., 2024; Kanamori et al., 2017). Additionally, we wanted to create a scale that allows users to capture more than one aspect of transprejudice despite being brief, thus enabling the measurement of more fine-grained information (Bergkvist & Rossiter, 2007). We conducted two studies to develop and validate the Brief Transprejudice Scale (BTS). In Study 1, we generated an initial item pool and conducted exploratory factor analysis (EFA) and Rasch modelling to identify a robust factor structure, and assessed internal reliability. In Study 2, we performed confirmatory factor analysis (CFA) to test the stability of this structure and examined concurrent, convergent and divergent validity by comparing the BTS with other relevant psychological measures.

Study 1

Method

Literature Search

We began by conducting an extensive search of scientific databases through the Web of Science citation index database. In addition to identifying existing validated measures of transprejudice, we used Web of Science to explore predictors of transprejudice to inform our scale development process. To identify a list of validated transprejudice scales, we used the following Boolean search terms: Transphobia AND (scale development), Transphobia AND (scale validation), Transprejudice AND (scale development), Transprejudice AND (scale validation). These search terms yielded 1,017 matches for possible inclusion.

We excluded any studies that used feeling thermometers (single-item scales) to operationalise transprejudice. We also excluded studies that did not explicitly develop and validate a measure of transprejudice for general use. Finally, we excluded studies in which the sole purpose was to validate an existing scale in a different cultural context – we focused instead on the original scales. Additionally, because our literature search uncovered an extensive systematic review of transprejudice scales (i.e., Morrison et al., 2017),

we inspected this review in case our own search had missed any transprejudice measures. These steps resulted in the 14 published scales highlighted in Table 1.

Item Generation

After conducting this exhaustive review of the literature, we generated a list of 90 scale items to assess attitudes towards transgender identity. We adapted some items from existing scales, for example, eight items were adapted from the Perceived Realistic Physical Threat Scale by Hellmann et al. (2022), who designed their scale to measure attitudes towards migrants. Because perceived threat has been found to be associated with prejudice against transgender people (transgender women in particular; see Hayes & Reiman, 2021), we included these items in the pool, replacing 'migrants' with 'transgender women/men'. We included one item verbatim from the T-KAB scale (i.e., 'Sex and gender are the same thing'; Clark & Hughto, 2020). We then generated an additional four items centred around sex and gender. This was done in an effort to capture participants' beliefs about gender identity versus biological sex. We also included some reworded items from the ATTMW scale (Billard, 2018a). The remaining items were novel, informed by both qualitative and quantitative research. For instance, Watson et al. (2019) conducted research identifying common themes of trans discrimination from the perspective of trans people, so we generated items to tap into such themes (e.g., 'Transgender people should have their gender identity respected by friends, family and coworkers (reverse-scored)', 'Refusing to work with someone because they are transgender is not discriminatory').

Next, we conducted three cycles of item scrutiny to identify which of the 90 items were potentially redundant, confusing or too vague. This resulted in a final pool of 52 items. This process was firmly driven by our review of the transprejudice literature, wherein we identified three main underlying conceptual domains. The first relates to perceptions about the legitimacy of trans women and men identifying and being treated as women/men (e.g., Billard, 2018a; Elischberger et al., 2016; White & Jenkins, 2017). Also reflected was a recurring perception that trans women in particular pose some form of threat or risk to the safety of cis women (Amery & Mondon, 2024; Atwood et al., 2024). Given the lack of evidence to support this (Morgenroth et al., 2024), it constitutes a form of transprejudice. As such, we retained 27 items from the initial item pool to capture attitudes related to trans identity (e.g., essentialist beliefs about gender/sex or gender identity), access of trans women and men to gendered spaces as well as misconceptions participants may have about risks to safety posed by trans women and men.

Our literature review also highlighted aspects of trans-related discrimination linked to legislation, employment practices and access to healthcare (Bauer et al., 2014; Beattie, 2022; Dowers et al., 2019). Given that transprejudice can sometimes take the form of support for policies or practices that disproportionately affect trans women and men (Billard, 2018a), we retained 16 items aimed at capturing this aspect of prejudice. These items aimed to capture policy-based attitudes such as those related to trans people receiving gender-affirming healthcare from the National Health Service (NHS), government recognition of self-declared gender identity and equal rights for cis and trans people.

Finally, the literature review also revealed a link between individuals' moral beliefs and their attitude towards trans people (e.g., Kay & Dimakis, 2024; Sayılan et al., 2020). As such, we took insights from moral foundations theory (Haidt & Joseph, 2008), which suggests that there are innate psychological systems that help shape a society's moral norms as well as individuals' emotional reactions to violations of those norms. Moral foundations theory outlines multiple mechanisms or *foundations* that underpin moral judgements, which can elicit either positive or negative affect as reaction to perceived moral violations. For instance, the *fairness moral foundation* might elicit positive trans attitudes, whereas the *purity moral foundation* might elicit negative trans

attitudes (Graham et al., 2009; Hadarics & Kende, 2018). Thus, we retained nine items intended to capture attitudes towards trans people that might be driven by moralism and/or negative affect.

Although most of the items in the initial pool of 90 were broadly within these three conceptual domains, we eliminated 38 that we deemed redundant or poorly worded upon closer scrutiny. Of the remaining 52 items, 20 were specific to transgender women, 19 were specific to transgender men[2] and 13 related to transgender people in general (see the Appendix for the list of items).

We note here that the development of this scale focused primarily on prejudice against trans women and men. This was partly driven by the extant literature, a larger proportion of which was centred around binary trans experiences, with relatively less research on nonbinary trans experiences. Although there are some items within our scale that refer to transgender people (rather than transgender women/men), they do not specifically refer to nonbinary people. This means that whilst this scale may capture prejudice against trans nonbinary people as well, it may not capture important nuances unique to the prejudice against nonbinary individuals. Attempting to create a scale that measures prejudices towards binary and nonbinary trans people may limit content validity, which is one of the core objectives of scale validation (Boateng et al., 2018; Morrison et al., 2017). Thus, we focused on the development of a scale measuring prejudice against trans women and men specifically.

Participants

We aimed to recruit a minimum of 300 participants, based on research recommending a sample size of 150–300 for EFA (Kyriazos, 2018; Watkins, 2018), plus an additional 15% to allow for attrition and pilot testing (we conducted a pilot test of the 52 items to estimate the length of the study). We recruited a total of 346 UK-based participants from the online recruitment platform Prolific. Five participants failed an embedded attention check, so their data were excluded from analysis. The attention check read: 'It is important to pay close attention. Please select "strongly disagree" below'. This left a final sample of 341 participants aged between 19 and 77 years ($M = 40.23$, $SD = 13.36$). The sample consisted of 216 cisgender women, 100 cisgender men, four nonbinary persons, 14 who described their gender using other terms and seven who did not report their gender identity. In relation to sexual orientation, 282 participants identified as heterosexual, 20 as bisexual, 16 as gay/lesbian, eight as pansexual, three as asexual, three as queer, one as gynosexual, one reported multiple sexual identities and seven did not reported their sexual identity. Regarding race/ethnicity, 286 participants identified as White/Caucasian, 16 as Black/African, 12 as South Asian, eight as East Asian, four as Southeast Asian, six as mixed race and nine did not report or specify their race/ethnicity. We also measured political orientation on a 5-point scale (1 = *extremely liberal*, 5 = *extremely conservative*); the mean score was 2.65 ($SD = 0.87$).

Materials

Scale Items. We presented the 52 items (see Appendix) to participants and asked them to indicate their agreement with each of them. Responses were given on a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*; with the midpoint labelled as *neither agree nor disagree*). Example items are 'A transgender woman is just another kind of woman' (reversed), 'Transgender men make me feel uneasy' and 'I think there is something immoral about transgender people'. Higher scores indicate higher levels of transprejudice.

Additional Measures. We also asked participants to indicate their prior knowledge about and contact with trans people. We measured prior knowledge and contact with the intention of controlling for these in our analysis because both are associated with lower levels of transprejudice (e.g., Cheso et al., 2024).[3]

Procedure

We preregistered this study on the Open Science Framework platform (OSF; https://osf.io/njwps/overview?view_only=77a33a06c6b34

fc8a4be5b4d991cc103). Participants accessed the Qualtrics-hosted online study through Prolific. They were informed that they would be asked questions about their views regarding transgender people and issues related to transgender identity. They were then presented with definitions of transgender and cisgender women and men: a transgender woman is someone whose sex assigned at birth was male but whose gender identity is woman, and a cisgender woman is someone whose sex assigned at birth was female and whose gender identity is woman. After reading the definitions, participants were presented with the 52 items, randomised across participants. They then completed their demographic information and transgender knowledge and contact measures before debriefing. Both this study and Study 2 received ethical clearance from Royal Holloway, University of London and adhered to British Psychological Society ethics principles.

Results

EFA

We first conducted tests to ensure the data were suitable for factor analysis. Sampling adequacy was assessed using the Kaiser–Meyer–Olkin (KMO) criterion test; this was 0.97, indicating suitability for factor analysis. Bartlett's test of sphericity, $\chi^2(1326) = 20444.85$, $p < .001$, indicated sufficient variance in the data. Having met these assumptions, we proceeded with the EFA.

First, we conducted a parallel analysis on the 52 items using the 'psych' package (Revelle, 2024) in R (Version 4.4.0) to determine the number of factors. This suggested a four-factor solution. We then ran an EFA with four factors to determine factor loadings, using oblique rotation (promax) to allow for correlations between factors. To ensure robust factor retention, we set the factor loading criterion to .50. At this stage, 13 items failed to load onto any factors. Thus, we reran a parallel analysis on the remaining 39 items, resulting in a three-factor solution. We then conducted another factor analysis on the 39 items, this time specifying three factors to be retained. We continued this process of running a parallel analysis, running an EFA and removing items with factor loadings below .50 iteratively until all items loaded onto at least one factor, in line with recommended practice (see Cabrera-Nguyen, 2010). After the sixth iteration, 26 items loaded onto three factors (for R markdown scripts, data and html output of this iterative process, see https://osf.io/njwps/overview?view_only=77a33a06c6b34fc8a4be5b4d991cc103).

Rasch Analysis

Taking insights from item response theory (Hambleton & Swaminathan, 1985), we sought to assess the performance of the remaining 26 items using Rasch analysis (Rasch, 1980). Specifically, we aimed to examine the difficulty of each scale item to see which were either too easy or too difficult to endorse (Zamora-Araya et al., 2018). Rasch analysis has previously been used as part of scale development and validation processes (e.g., in validating an Italian version of the Persecutory Ideation Questionnaire, Colledani et al., 2024; to assess the psychometric properties of the Benevolent Sexism Subscale, Zamora-Araya et al., 2018). We conducted the Rasch analysis using the rating scale model – employed for analysis of Likert scale data. This enabled us to assess the measurement error for each item and identify any misfitting items based on two fit statistics: information-weighted fit (INFIT) and outlier-sensitive fit (OUTFIT; Linacre, 2002). These fit indices are mean-square values that help identify scale items with measurement variances that are significantly higher or lower than those estimated by the Rasch model. INFIT and OUTFIT mean-square values that fall within 0.5–1.5 are considered suitable for measurement (Linacre, 2002). Mean-square values that fall outside of this threshold are indicative of poor fit; such items might, for example, be worded so ambiguously as to elicit an erratic pattern of responses from participants (Zou & Bolt, 2023).

We conducted the Rasch analysis using Jamovi software (The Jamovi Project, 2024, version 2.5.6). Because the rating scale model (or polytomous model) in Rasch analysis requires the

minimum score of scale items to be set to zero, we recoded the items such that scores ranged from 0 (*strongly disagree*) to 6 (*strongly agree*). We then assessed the INFIT and OUTFIT (Linacre, 2002) mean-square values, first separately for each of the three factors and then entering all items into the model at the same time. Following this analysis, eight items in total were identified as falling outside the range of our INFIT–OUTFIT exclusion criterion of 0.5–1.5. A visual inspection of the misfitting items showed that many of them were centred around sex and gender identity (e.g., 'Sex is more important than gender identity', 'Participation in female-only sports should be based on sex and not gender identity'). This suggests that the distinction between sex and gender identity might not have been clear-cut, at least in the case of our sample. Another possibility is that the wording of the misfitting items was too ambiguous or confusing. We decided to eliminate these eight items from the factor model, thus leaving 18 in (see Table 2 for INFIT and OUTFIT mean-square values of the 26 scale items).

Given this reduction in scale items, we conducted another iteration of parallel and factor analyses on the remaining 18 items to identify changes in the factor model. A three-factor solution was again suggested by the parallel analysis; however, two scale items failed to load onto any factors. The two items were eliminated, leaving 16 loading onto a three-factor solution.

Assessing Model Fit

To ensure the factor model was a good fit for the data, we assessed the fit indices via a CFA. Most of the fit indices were a good fit for the data: CFI = .91, TLI = 0.89, SRMR = .06. However, the RMSEA index fell outside the acceptable range of fit, RMSEA = .13. Thus, we assessed the modification indices to identify aspects of the model that could be altered to improve model fit. This indicated that allowing one of the items to cross-load onto two factors would improve model fit (the item reads: 'Transgender women make me feel uneasy'). However, as the aim of scale development, typically, is to avoid items cross-loading onto multiple factors, we chose to eliminate this item instead. Additionally, modification indices indicated that one of the items should load onto a different factor (i.e., 'I find it upsetting that transgender men identify as men'). We reanalysed the model based on this change; however, this did not improve model fit. Thus, we eliminated this item as well. This left us with 14 scale items. Finally, the modification indices also indicated that allowing some of the items within each factor to covary would improve model fit. This included items that were mirrors of each other (e.g., 'A transgender man will never truly be a man', 'A transgender woman will never truly be a woman'). Implementing these changes to the factor model improved model fit, CFI = .97, TLI = 0.96, RMSEA = .08, SRMR = .04.

Factor Structure and Reliability

After conducting our factor and Rasch analyses, 14 items were retained in a three-factor solution.[4] Inspection of the items retained in each factor indicated the emergence of identifiable themes/subscales across the three dimensions: negative affect (five items), category exclusion (five items) and moralism (four items; see Table 3). This three-factor model explained 79% of the variance in the data, with negative affect accounting for 25%, category exclusion accounting for 30% and moralism accounting for 24%.

Next, we conducted reliability analyses to assess internal consistency. We calculated both Cronbach's alpha (α) and person separation reliability (*R*). Cronbach's alpha was calculated using the 'psych' package in R (Revelle, 2024), whilst person separation reliability was calculated using Jamovi software alongside Rasch analysis. Separation reliability is similar to Cronbach's alpha; a low value (i.e., $< .70$) suggests that the scale might not be sensitive enough to detect differing levels of transprejudice across participants (Brown & Bonsaksen, 2019; Robinson et al., 2019). Person separation reliability has been shown to be more conservative than Cronbach's

Table 2

Item Statistics for the Rasch Rating Scale Analysis Showing Measure, INFIT, and OUTFIT Values for the 26 Scale Items: Study 1

Item	Measure	INFIT	OUTFIT
Transgender women should be excluded from women-only spaces	0.07	1.21	1.13
A transgender woman will never truly be a woman	−0.05	1.12	1.01
I would struggle to think of transgender women as women	0.36	0.98	0.85
Transgender women make me feel uneasy	1.09	0.76	0.69
Transgender women scare me	1.60	1.08	0.87
Transgender women make me feel unsafe	1.43	1.07	0.95
Transgender men should be excluded from men-only spaces	0.52	1.23	1.14
A transgender man will never truly be a man	0.07	1.11	0.99
I find it upsetting that transgender men identify as men	1.52	0.77	0.70
Transgender men make me feel uneasy	1.39	0.90	0.86
Transgender men scare me	1.64	1.13	1.02
Transgender men make me feel angry	1.20	0.82	0.75
Transgender men make me feel unsafe	1.76	0.97	0.94
It should be against the law to be transgender	2.24	0.98	0.72
I do not trust transgender people	1.70	0.69	0.63
I think there is something immoral about transgender people	1.66	0.83	0.65
Transgender people are ruining the natural order of things in our society	1.18	0.98	0.86
Being transgender is wrong because it is unnatural	1.46	1.03	0.95
Sex is more important than gender identity	−0.29	1.25	1.61
Transgender men should only have access to gender-neutral spaces	0.19	1.48	1.88
Participation in male-only sports should be based on sex and not gender identity	−1.13	1.40	1.46
Compared to cisgender women, transgender women are innately masculine	0.28	1.24	1.46
Transgender women should be protected under the law as women (reverse-scored)	0.25	1.10	1.48
People should be allowed access to female-only or male-only spaces based on sex and not gender identity	−0.56	1.27	1.46
Participation in female-only sports should be based on sex and not gender identity	−1.39	1.71	2.13

Note. Items in bold were subsequently excluded due to INFIT/OUTFIT mean-square values falling outside of the 0.5−1.5 range.

Table 3

Scale Items, Means, Standard Deviations and Factor Loadings: Study 1

Items	*M* (*SD*)	Factor loadings		
		Negative affect	Category exclusion	Moralism
Transgender men scare me	2.11 (1.34)	.88		
Transgender women scare me	2.40 (1.63)	.87		
Transgender men make me feel unsafe	2.52 (1.50)	.85		
Transgender women make me feel unsafe	2.55 (1.72)	.66		
Transgender men make me feel uneasy	2.60 (2.11)	.65		
A transgender man will never truly be a man	4.06 (2.11)		.99	
A transgender woman will never truly be a woman	4.19 (2.10)		.97	
Transgender women should be excluded from women-only spaces	4.05 (2.09)		.90	
I would struggle to think of transgender women as women	3.72 (2.08)		.78	
Transgender men should be excluded from men-only spaces	3.54 (1.72)		.73	
I think there is something immoral about transgender people	2.33 (1.66)			.93
Being transgender is wrong because it is unnatural	2.53 (1.86)			.89
It should be against the law to be transgender	1.87 (1.38)			.87
Transgender people are ruining the natural order of things in our society	2.81 (1.85)			.65

alpha (Anselmi et al., 2019). Thus, its inclusion in our analyses is aimed at increasing the robustness of our reliability checks. All three subscales showed good reliability: negative affect, $\alpha = .93$, $R = .78$; category exclusion, $\alpha = .95$, $R = .89$; moralism, $\alpha = .95$; $R = .80$. The overall reliability of the 14-item scale was also excellent, $\alpha = .96$, $R = .91$.

Discussion

In this study, we present a new measure of transprejudice. Through an iterative process of EFA and Rasch analysis, we developed a 14-item scale containing three subscales: negative affect, category exclusion and moralism. The overall scale, as well as the three subscales, showed excellent internal consistency. This study represents the first step in the process of scale validation.

Study 2

Having established a three-factor solution in Study 1, the aim of Study 2 was to test the external validity of the newly developed BTS. We conducted CFA to test whether the factor structure would reemerge. We also tested how our newly developed scale related to other psychometric measures. This was done by testing the concurrent, convergent and divergent validity of the BTS.

Concurrent validity is a form of criterion-related validity that tests whether a new scale captures the same construct as another scale. Concurrent validity is established when there is a significant correlation (in the expected direction) between a newly developed measure and a previously developed one of the same construct (Boateng et al., 2018; Morrison et al., 2017). We used two scales to test the concurrent validity of

our measure: the revised short form of the Genderism and Transphobia Scale (GTS-R-SF; Tebbe et al., 2014) and the ATTMW scale (Billard, 2018a). We used the GTS-R-SF because various versions of the GTS (Hill & Willoughby, 2005) have been used to measure transprejudice in over two dozen published studies (e.g., Cabeldue et al., 2018; Chen & Anderson, 2016; Costa & Davies, 2012; see also Billard, 2018a). Additionally, the GTS was highlighted by Morrison et al. (2017) as a near 'gold standard' transprejudice scale in their systematic review (but see our discussion of its limitations on pp. 2 and in Table 1). The ATTMW scale was included as it is a more recent measure of transprejudice. To show concurrent validity, we expected a significant positive correlation between these scales and the BTS.

Convergent validity is a form of construct validity used to test whether a new scale is measuring a construct that is related to other psychological constructs (Morrison et al., 2017). To establish convergent validity, there must be a significant correlation between a newly developed scale and a scale measuring a construct that is theoretically or empirically related yet distinct from the construct of interest (Boateng et al., 2018). For example, transprejudice has been empirically linked to other psychological constructs such as gender essentialism, religiosity, RWA and SDO (Campbell et al., 2019; Hatch et al., 2022; Makwana et al., 2018). Thus, we sought to test the convergent validity of our scale using measures of RWA and SDO. We anticipated a significant positive correlation in both cases.

Finally, divergent validity is indicated by a weak or nonsignificant correlation between the scale of interest and scales measuring unrelated constructs (Boateng et al., 2018; Morrison et al., 2017). We tested the divergent validity of the BTS using the Use of Creative Cognition Scale (Rogaten & Moneta, 2015) and the Brief Resilience Scale (Smith et al., 2008). The first scale measures one's creativity, whilst the second measures one's ability to deal with difficult situations. We expected there to be weak (i.e., $r \leqslant .20$) or nonsignificant correlations between these and the BTS. Study 2 was preregistered on the OSF (see https://doi.org/10.17605/OSF.IO/BMKYJ).

Method

Participants

To determine sample size for this study, we implemented previously recommended heuristics for CFA studies. Specifically, we aimed for a minimum of 10:*q* participants, where *q* is the number of parameters in the factor model (see Koran, 2016). We estimated the number of parameters in R using the 'lavaan' package (Rosseel, 2012). Based on a three-factor solution with 14 indicator variables, the number of parameters was estimated to be 31; thus, we aimed to recruit a minimum of 310 participants. To ensure that our study was not underpowered and to account for any attrition, we recruited a representative U.K. sample of 391 participants from Prolific. Participants were aged between 18 and 87 years ($M = 46.23$, $SD = 15.65$). The sample was comprised of 186 individuals that identified as cisgender women, 164 as cisgender men, three as nonbinary, two as transgender men, 22 described their gender using other terms and 14 did not report gender identity. In relation to sexual identity, 343 identified as heterosexual, 17 as bisexual, 12 as gay/lesbian, two reported multiple sexual identities, one identified as pansexual, four as asexual, one as queer and 11 did not report or specify sexual identity. Regarding race/ethnicity, 327 participants identified as White/Caucasian, 17 as South Asian, 11 as Black/African, six as East Asian, seven as Southeast Asian, four as mixed race, three as Hispanic/Latinx, two as Middle Eastern and 14 did not report or specify race/ethnicity. As in Study 1, we measured political orientation on a 5-point scale ($M = 2.75$, $SD = 0.98$).

Materials

Responses for all scale items were given on a 7-point scale (1 = *strongly disagree*, 7 = *strongly*

agree, with the midpoint labelled as *neither agree nor disagree*).

The BTS. Participants were presented with the final 14 items that emerged from the first study. We also reintroduced two items that were eliminated from Study 1 after inspection of modification indices (i.e., 'Transgender women make me feel uneasy', 'I find it upsetting that transgender men identify as men'). Because both items had a factor loading above .50, we wanted to confirm whether they would again contribute to misfit of the factor model or whether it would be appropriate to keep them as part of the scale. However, as in Study 1, our analyses showed that both items worsened the fit indices of the factor model, and thus they were subsequently dropped. Higher scores on the BTS indicate higher levels of transprejudice.

GTS-R-SF. To test concurrent validity, we used the short and revised version of the GTS (Tebbe et al., 2014). The scale contains 13 items (e.g., 'I would avoid talking to a woman if I knew she had a surgically created penis and testicles'). This scale showed excellent reliability, $\alpha = .92$. Higher scores indicate higher levels of transprejudice.

ATTMW Scale. As a second test of concurrent validity, we used the 24-item ATTMW Scale (e.g., 'Transgender women are only able to look like women but not be women'; Billard, 2018a). This scale also showed excellent reliability, $\alpha = .99$. Higher scores indicate higher levels of transprejudice.

Short Social Dominance Orientation Scale. To test convergent validity, we used the four-item Short Social Dominance Orientation Scale (e.g., 'Superior groups should dominate inferior groups'; Pratto et al., 2012). Higher scores indicate higher levels of SDO. This scale showed good internal consistency, $\alpha = .81$.

Very Short Authoritarianism Scale. We used the six-item Very Short Authoritarianism Scale to measure RWA (Bizumic & Duckitt, 2018) as a second test of convergent validity. An example item reads, 'What our country needs most is discipline, with everyone following our leaders in unity'. Higher scores indicate higher levels of RWA. The internal consistency of this scale was acceptable, $\alpha = .78$.

Use of Creative Cognition Scale. To test the divergent validity of our scale, we presented participants with the five-item Use of Creative Cognition Scale (e.g., 'I find effective solutions by combining multiple ideas'; Rogaten & Moneta, 2015). Higher scores indicate higher levels of creative cognition. This scale showed good internal consistency, $\alpha = .83$.

The Brief Resilience Scale. As our second test of divergent validity, we used the six-item Brief Resilience Scale (e.g., 'It does not take me long to recover from a stressful event'; Smith et al., 2008). Higher scores indicate higher levels of personal resilience. This scale showed excellent internal consistency, $\alpha = .91$.

Procedure

This online study was hosted on Qualtrics. As in Study 1, participants were presented with definitions of transgender and cisgender women and men before responding to the BTS. After this, they were presented, in random order, with the six scales outlined in the Measures section above (for item order within each scale, see Study 2, Measures section at https://osf.io/njwps/overview?view_only=77a33a06c6b34fc8a4be5b4d991cc103). Participants then completed their demographic information, (i.e., age, gender identity, sexual identity, political orientation and ethnicity) before being debriefed.

Results

EFA

We first conducted preliminary tests to ensure that the assumptions for factor analysis were met. These tests showed that the data were indeed

suitable for factor analysis: KMO criterion test, 0.92; Bartlett's test of sphericity, $\chi^2(91) = 7412.31$, $p < .001$.

We conducted a parallel analysis on the 16 items to determine the number of factors suggested for a factor solution. In alignment with the results of the first study, the parallel analysis suggested a three-factor solution. We then ran an EFA with three factors using oblique rotation (promax) and factor loading set to .50. One of the items ('I find it upsetting that transgender men identify as men') failed to load onto any factors; this item was dropped from the model, resulting in a 15-item model.

CFA

Assessing Model Fit. We conducted a CFA on the 15 items to assess the fit of the three-factor model: CFI = .91, TLI = 0.89, SRMR = .05, RMSEA = .15. Because the RMSEA fit index fell outside the acceptable range of goodness of fit, we assessed the modification indices to identify means of improving model fit. Modification indices with the highest values are most likely to improve overall model fit (Hooper et al., 2008; Kyriazos, 2018), we therefore set this criterion to ⩾ 20. The modification indices indicated that allowing one of the items to cross-load onto all three factors would improve model fit (the item in question was 'Transgender women make me feel uneasy'). However, as the aim of scale development is typically to avoid items cross-loading onto multiple factors, we chose to eliminate this item instead. Thus, the two items that had been reintroduced from Study 1 were dropped. The modification indices also indicated allowing covariance between equivalent items within the same factor (e.g., 'Transgender women scare me' was allowed to covary with 'Transgender men scare me'). Additionally, the modification indices suggested allowing covariance between some of the scale items belonging to the same factor. Specifically, 'Transgender men scare me' was allowed to covary with 'Transgender men make me feel unsafe' and 'Transgender men make me feel uneasy', while 'Transgender women scare me' was allowed to covary with 'Transgender women make me feel unsafe'. There is precedent for allowing between-item measurement error correlations in scale validation, to account for potential systematic measurement bias (Goretzko et al., 2024; Montoya & Edwards, 2021).

Implementing these changes improved the model fit: CFI = .98, TLI = 0.98, RMSEA = .07, 90% CI_{RMSEA} [0.06, 0.08], SRMR = .03. Thus, a three-factor solution was confirmed, with 14 items in total (see Table 4 for factor loadings, means and standard deviations for the 14-items).[5] As in Study 1, we analysed both Cronbach's alpha (α) and person separation reliability (R) to assess the reliability of our scale (note that we did not preregister the person separation reliability analysis for Study 2; however, we chose to include this analysis to remain consistent with the first study). All three subscales showed excellent internal consistency: negative affect, $\alpha = .95$, $R = .87$; category exclusion, $\alpha = .96$, $R = .91$; moralism, $\alpha = .96$, $R = .81$. The overall scale also showed excellent internal consistency, $\alpha = .97$, $R = .92$.

Criterion-Related and Construct Validity

After confirming our three-factor model, we conducted zero-order correlations to test the concurrent, convergent and divergent validity of our scale (see Table 5 for correlations, including subscales).

Concurrent Validity

We found a significant positive correlation between our BTS and the ATTMW and GTS-R-SF scales (see Table 5). This finding indicates that the BTS captures transprejudice as intended (Morrison et al., 2017) and provides evidence of concurrent validity.

Convergent Validity

As predicted, we found a significant positive correlation between our transprejudice measure and both the Short Social Dominance Orientation Scale and the Very Short Authoritarianism Scale. Notably, the correlation coefficients here were

Table 4

Factor Loadings, Means and Standard Deviations of the 14 Scale Items: Study 2

Items	*M* (*SD*)	Factor loadings		
		Negative affect	Category exclusion	Moralism
Transgender women scare me	2.46 (1.55)	.91		
Transgender women make me feel unsafe	2.45 (1.60)	.92		
Transgender men scare me	2.42 (1.52)	.89		
Transgender men make me feel unsafe	2.33 (1.47)	.88		
Transgender men make me feel uneasy	2.70 (1.70)	.96		
A transgender woman will never truly be a woman	4.20 (2.01)		.92	
Transgender women should be excluded from women-only spaces	4.00 (2.00)		.88	
I would struggle to think of transgender women as women	3.84 (1.97)		.94	
A transgender man will never truly be a man	4.27 (2.04)		.91	
Transgender men should be excluded from men-only spaces	3.74 (1.97)		.87	
I think there is something immoral about transgender people	2.42 (1.74)			.93
It should be against the law to be transgender	2.00 (1.51)			.90
Being transgender is wrong because it is unnatural	2.62 (1.82)			.95
Transgender people are ruining the natural order of things in our society	2.64 (1.88)			.92

smaller than the correlations between the transprejudice scales (see Table 5). This indicates that our newly developed scale is associated with distinct but conceptually linked constructs (here, SDO and RWA), thus providing support for convergent validity (Boateng et al., 2018; Morrison et al., 2017).

Divergent Validity

Finally, there was no significant correlation between our transprejudice measure and either the Use of Creative Cognition Scale or the Brief Resilience Scale. This provides support for divergent validity as it indicates that the BTS is not associated with constructs unrelated to transprejudice. It is worth noting that divergent and discriminant validity were also shown via the BTS correlations with SDO and RWA ($r = .55$ and $r = .59$, respectively). Decades of research on intergroup attitudes and prejudice predicts that SDO and RWA should be moderate-to-strong positive antecedents of prejudice toward stigmatised groups (Hatch et al., 2022; Perez-Arche & Miller, 2021). Thus, the BTS shows correlations with these antecedents, in the predicted direction, at magnitudes consistent with the broader prejudice literature. This indicates the BTS is embedded in the expected theoretical network of prejudice.

Multiple Group Factor Analyses: Testing Measurement Invariance

Measurement invariance is used to determine whether a scale elicits consistent responses across different samples and subpopulations (Wang et al., 2018). Thus, measurement invariance analyses add to the robustness of the scale development process by demonstrating the generalisability

Table 5

Zero-Order Correlations Showing Concurrent, Convergent and Divergent Validity: Study 2

	M (*SD*)	1	2	3	4	5	6	7	8	9
1. Brief Transprejudice Scale	3.01 (1.48)									
2. Negative Affect Subscale	2.47 (1.45)	.89***								
3. Category Exclusion Subscale	4.01 (1.86)	.90***	.67***							
4. Moralism Subscale	2.42 (1.64)	.90***	.77***	.70***						
5. ATTMW	3.69 (1.80)	.91***	.71***	.92***	.80***					
6. GTS-R-SF	2.28 (1.12)	.89***	.78***	.74***	.90***	.82***				
7. SDO	2.60 (1.13)	.55***	.48***	.52***	.47***	.53***	.51***			
8. RWA	3.40 (1.12)	.59***	.48***	.49***	.62***	.57***	.61***	.36***		
9. Creativity	5.43 (0.75)	−.01	−.02	−.01	.02	−.02	.01	−.08	.12*	
10. Resilience	4.44 (1.18)	.03	−.03	.07	.02	.06	−.01	.05	.02	.28***

Note. ATTMW = Attitudes Towards Transgender Men and Women Scale; GTS-R-SF = Genderism and Transphobia Scale Revised Short Form. SDO = social dominance orientation measured via the Short Social Dominance Orientation Scale; RWA = right-wing authoritarianism measured using the Very Short Authoritarianism Scale. Creativity was measured via the Use of Creative Cognition Scale, and resilience using The Brief Resilience Scale.
*$p < .05$. ***$p < .001$.

of the scale. We checked for four types of invariance for our factor solution: (a) dimensional invariance, shown if the scale items load onto the factors as expected across the subsamples being tested; (b) configural invariance, shown if the grouped model has an acceptable fit; (c) metric invariance, tested by constraining factor loadings to be equal across groups and comparing the fit indices of the nested metric model to the nested configural model; (d) scalar invariance, tested by constraining both the size of factor loadings and item intercepts to be equal across groups, and comparing the fit indices to the nested metric and configural models (for review of measurement invariance methods, see Putnick & Bornstein, 2016). There have been recommendations to use alternative fit indices (rather than the chi-square statistic) to assess measurement invariance because chi-square tends to be overly sensitive to sample size (Goretzko et al., 2024; Hooper et al., 2008; Wang et al., 2018). Thus, we followed these recommendations. Specifically, difference in fit indices between the configural, metric and scalar models should be no higher than CFI = −.01, RMSEA = .02, metric invariance =.03 and scalar invariance of SRMR = .02 (Fischer & Karl, 2019; Putnick & Bornstein, 2016). To perform these four measurement invariance tests, we conducted multiple confirmatory group factor analyses between Study 1 and Study 2, based on participants' gender and political orientation. We chose these two variables because extant literature has shown that men tend to score higher on transprejudice measures compared to women (Billard, 2018a). Additionally, conservatives tend to score higher that liberals on transprejudice measures (e.g., Hatch et al., 2022). Indeed, this was the case in both our studies (see Tables A1 and A2 in the Appendix showing descriptive and inferential statistics for Studies 1 and 2, grouped by gender[6] and political orientation, respectively).

Gender. We pooled the data from Studies 1 and 2 (N = 732), then grouped participants by gender (i.e., cisgender women: n = 402; cisgender men: n = 264), excluding data from noncisgender participants. We first ran EFA separately for each gender group to ensure that the 14 items loaded onto the same three factors as expected. This was indeed the case, so dimensional invariance was achieved. We went on to test for the remaining three types of measurement invariance using the 'lavaan' package in R (Rosseel, 2012). To check for configural invariance, we inspected the fit indices for the three-factor model, which showed all were within acceptable range (see Table 6 for fit indices). To test metric invariance, we constrained the factor loadings in the CFA model to be equal across gender groups. To test for scalar invariance, we constrained both the factor loadings and item intercepts to be equal across gender groups. We then used the 'compareFit' function from the 'semTools' package (Jorgensen et al., 2022) to test whether there were significant differences in model fit as a result of these constraints. We found no significant difference in the metric versus scalar CFA models, which suggests measurement invariance (see Table 6 for comparative fit indices). Thus, the BTS appeared to perform consistently across Studies 1 and 2 for cisgender women and men.

Political Orientation. We created three grouping variables for political orientation: liberal, moderate and conservative. Participants who indicated that they were '*extremely liberal*' or '*somewhat liberal*' were assigned to the liberal group (n = 303); those who indicated they were '*moderate*' were assigned to the moderate group (n = 284); and those who indicated they were '*somewhat conservative*' or '*extremely conservative*' were assigned to the conservative group (n = 145). We then ran EFA separately on the three political groups to test for dimensional invariance. The 14 items loaded onto the three factors as expected; thus, dimensional invariance was achieved for all three subgroups of political orientation. We then tested configural, metric and scalar invariance. As with gender (above), we found that all fit indices fell within acceptable range (see Table 6), and there was no significant difference in fit indices between the nested models. Thus, the BTS

Table 6

Comparative Fit Indices Showing Configural, Metric and Scalar Invariance for Gender and Political Orientation Groups

CFA model	CFI	RMSEA [90% CI]	SRMR	Model comparison	ΔCFI	ΔRMSEA	ΔSRMR
Gender							
Configural	.97	.08 [0.07, 0.09]	.04	–	–	–	–
Metric	.97	.08 [0.06, 0.09]	.04	Configural	.000	−.003	.000
Scalar	.96	.08 [0.07, 0.09]	.04	Metric	−.003	.002	.002
Political orientation							
Configural	.97	.08 [0.07, 0.09]	.05	–	–	–	–
Metric	.97	.08 [0.07, 0.09]	.06	Configural	−.002	−.002	.010
Scalar	.97	.08 [0.07, 0.09]	.06	Metric	−.002	−.001	.000

demonstrated measurement invariance across political orientation − across both studies, liberals, moderates and conservatives responded to the scale items consistently.

General Discussion

We present a new psychometric measure of transprejudice developed and validated across two studies. In the first study, we established a three-factor solution through a series of EFAs as well as Rasch analysis. In the second study, we verified this three-factor model through CFA. Our scale also demonstrated good internal consistency and validity, indicating that it is a psychometrically sound measure. The development and validation of the 14-item BTS represents an important contribution to research on trans-related attitudes.

To our knowledge, this is the shortest multidimensional scale measuring transprejudice. This will allow researchers to capture the complexities of transprejudice in a way that a unidimensional scale might not. Indeed, the literature on transprejudice illustrates that trans-related attitudes tend to be multifaceted, from categorical beliefs about gender/sex to rejection of gender nonconformity (Hatch et al., 2022; Makwana et al., 2018). The BTS is a measure short enough to limit participant cognitive overload yet able to capture multiple aspects of transprejudice. For researchers looking to explore how different constructs relate to this multifaceted scale, we recommend reporting and analysing the three BTS subscale scores (negative affect, category exclusion and moralism) separately; for example, entering them as separate predictors in a regression. This approach preserves the conceptual gains of a multidimensional measure: it allows researchers to detect, for instance, whether a given construct (e.g., RWA, SDO, religiosity) relates differently to affective, categorical and moralistic forms of transprejudice – distinctions that a single composite score would mask. Below, we provide an overview of the three subscales identified in the BTS.

We labelled the negative affect dimension as such because the items contained therein highlight an aspect of transprejudice associated with having a negative affective reaction towards trans individuals (e.g., 'Transgender women/men scare me'). Given that transprejudice/transphobia is defined as a fear or aversion towards people who are trans (Tebbe et al., 2014), it is unsurprising that this subscale emerged. Neither the GTS-R-SF (Tebbe et al., 2014) nor the ATTMW (Billard, 2018a) scales contain items that focus specifically on emotional aversion. Moreover, compared to the other subscales, there is a weaker correlation between the Negative Affect Subscale

and both the GTS-R-SF and ATTMW scales. This suggests that the Negative Affect Subscale captures an aspect of transprejudice that is perhaps not captured by these previous measures of transprejudice.

The Category Exclusion Subscale contains items that reflect a tendency to adhere to a strict categorisation of gender group membership (e.g., 'A transgender woman/man will never truly be a woman/man'). This is a dimension of transprejudice that has emerged in previous scale validations, albeit under a different label. For example, the sex/gender beliefs factor in the Transgender Attitudes and Beliefs Scale (TABS; Kanamori et al., 2017; see also Jones et al., 2024) has items such as 'If you are assigned male at birth, nothing you do will change that'. Indeed, many of the items in the ATTMW scale (Billard, 2018a) seem to be related to beliefs about gender categorisation (e.g., 'Transgender women are only able to look like women but not be women'). The correlations in Table 5 support this assertion given that the ATTMW scale is most highly correlated with the Category Exclusion Subscale.

The final dimension of the BTS is the Moralism Subscale, labelled as such because these items reflect attitudes towards trans people that are driven by beliefs related to morality (e.g., 'I think there is something immoral about transgender people', 'It should be against the law to be transgender'). These items appear to reflect the *authority* and *purity foundations* of the moral foundations theory (Graham et al., 2009). These foundations are classed as *binding foundations*, that is, a domain of morality that acts as a community-protecting mechanism whereby people are assigned to groups with distinct roles (Graham et al., 2011). Anyone violating these roles is perceived to be harming the larger community and judged as immoral (Hadarics & Kende, 2018). Interestingly, Table 5 shows that the GTS-R-SF scale correlates the highest with the Moralism Subscale. Many of the items in the GTS-R-SF do seem to reflect attitudes driven by some moral code (e.g., 'A man who dresses as a woman is a pervert', 'Sex change operations are morally wrong'). The retention of the moralism factor across both our studies suggests that transprejudice is driven, at least in part, by what people perceive to be moral versus immoral behaviour. This is an avenue that has received little attention in the transprejudice literature and is worth further exploration.

Our scale builds on extant transprejudice measures by utilising insights from item response theory, specifically using Rasch modelling. By conducting Rasch analyses, we were able to identify and eliminate items that elicited an erratic pattern of responses from participants. Rasch analysis has been used in scale development and validation. However, to our knowledge, no other research has implemented this analysis in the development of a transprejudice measure. By conducting Rasch analysis as part of our scale development, we have added to the reliability and utility of our measure.

Another way in which the BTS builds on extant transprejudice measures is by demonstrating its measurement invariance across subgroups of gender and political orientation. Our findings show that participants belonging to different gender and/or political groups interpreted and responded to the BTS in a consistent manner. The finding of measurement invariance across both studies demonstrates the generalisability of the BTS.

Limitations

The BTS provides researchers with an updated and psychometrically sound measure of transprejudice. However, our research is not without limitations. For instance, we did not obtain input directly from trans individuals. Although we did an extensive review of extant literature on transprejudice (both quantitative and qualitative), we did not directly seek input from trans people about the items we generated in the development phase. In their systematic review, Morrison et al. (2017) emphasise the importance of including transgender people in the development and validation process of transprejudice scales. Doing so may have highlighted aspects of transprejudice that are less obvious to individuals who are not trans. That said, scale validation is an ongoing process that should extend beyond the initial

developers of a measure (Flake et al., 2017). Future research might consider a more rigorous assessment of the content validity of this scale, for instance, by subjecting the items to the scrutiny of focus group discussions led by transgender individuals.

Another limitation is that none of the items in the final version of the scale are negatively worded, that is, none of the items are reverse-scored. It is typically recommended that scales be comprised of both positively and negatively worded items to reduce acquiescence bias (D. Watson, 1992). However, having a mix of positively and negatively worded items might not minimise acquiescence bias as effectively as once thought. Chyung et al. (2018) argue that, linguistically speaking, somewhat disagreeing with a statement is not necessarily an exact opposite of somewhat agreeing with said statement. In the context of our scale, a participant might somewhat agree with the statement 'I would struggle to think of transgender women as women', but they might disagree or even strongly disagree with a negatively worded equivalent, for example, 'I would not struggle to think of transgender women as women' – indicating a higher score of prejudice in the latter statement compared to the former. Chyung et al. (2018) argue that using a mix of positive and negative statements may actually compromise the internal validity of a scale.

Despite these limitations, this work offers a useful contribution to extant research on trans-related attitudes. We have implemented seldomly used analysis techniques (e.g., Rasch analysis, person separation reliability) in an effort to add to the rigour and robustness of our scale validation. The BTS has been validated on a U.K. representative sample and has shown strong internal consistency and validity. Therefore, this 14-item multifactor scale offers a useful psychometric tool that can be used to advance research on transprejudice.

ORCID iDs

Diana Cheso https://orcid.org/0009-0007-6293-0843
R. Thora Bjornsdottir https://orcid.org/0000-0002-1016-3829
Ryan McKay https://orcid.org/0000-0001-7781-1539
Hanna Zagefka https://orcid.org/0000-0003-1598-0059

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Notes

1. In line with wider social psychology literature, we have chosen to use the term transprejudice over transphobia because it more accurately encapsulates the nature of this attitude as one that consists of affective, cognitive and behavioural tendencies (Ching, 2022). Additionally, transphobia was initially defined as an emotional disgust towards trans people (Hill & Willoughby, 2005), a narrower interpretation of this attitude.
2. There were only 19 items related to transgender men because the following item was specific to trans women: 'Transgender women threaten the safety and well-being of cisgender women'. Our literature review suggested that perceptions of threats to safety tend to be evoked by trans women more than trans men.
3. Although this was part of our preregistration, we subsequently opted not to include contact and knowledge in our analysis. This was done to maintain the generalisability of the scale, given that its purpose is to capture both higher and lower levels of prejudice.
4. We excluded the data of the few noncisgender participants and reran the EFA to check whether this factor solution would change. As the pattern remained the same when data from these participants were excluded, we chose to include them for the sake of statistical power.
5. As with Study 1, we ran factor analyses excluding data from noncisgender participants and found no significant differences in the factor structure or fit indices. As such, we chose to retain their data.

6. For study 2, the difference in BTS scores between women and men was only marginally significant, $p = .065$

References

Acker, G. M. (2017). Transphobia among students majoring in the helping professions. *Journal of Homosexuality*, *64*(14), 2011–2029. https://doi.org/10.1080/00918369.2017.1293404

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science: A call to action. *European Journal of Psychological Assessment*, *38*(1), 1–5. https://doi.org/10.1027/1015-5759/a000699

Amery, F., & Mondon, A. (2024). Othering, peaking, populism and moral panics: The reactionary strategies of organised transphobia. *The Sociological Review*, *73*(3), 1–17. https://doi.org/10.1177/00380261241242283

Amigo-Ventureira, A. M., Durán, M., & DePalma, R. (2025). Exploring the relationship between transphobia and homophobia and other demographic factors among practicing and future primary school teachers. *Research in Education*, *121*(1), 19–36. https://doi.org/10.1177/00345237231163048

Amnesty International. (2023, May 11). *UK: UN view on trans rights is 'much-needed common sense'*. https://www.amnesty.org.uk/press-releases/uk-un-view-trans-rights-much-needed-common-sense

Anselmi, P., Colledani, D., & Robusto, E. (2019). A comparison of classical and modern measures of internal consistency. *Frontiers in Psychology*, *10*, Article 2714. https://doi.org/10.3389/fpsyg.2019.02714

Atwood, S., Morgenroth, T., & Olson, K. R. (2024). Gender essentialism and benevolent sexism in anti-trans rhetoric. *Social Issues and Policy Review*, *18*(1), 171–193. https://doi.org/10.1111/sipr.12099

Bauer, G. R., Scheim, A. I., Deutsch, M. B., & Massarella, C. (2014). Reported emergency department avoidance, use, and experiences of transgender persons in Ontario, Canada: Results from a respondent-driven sampling survey. *Annals of Emergency Medicine*, *63*(6), 713–720. https://doi.org/10.1016/j.annemergmed.2013.09.027

Beattie, C. (2022). High court should not restrict access to puberty blockers for minors. *Journal of Medical Ethics*, *48*(1), 71–76. https://doi.org/10.1136/medethics-2020-107055

Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research*, *44*(2), 175–184. https://doi.org/10.1509/jmkr.44.2.175

Billard, T. J. (2018a). Attitudes toward transgender men and women: Development and validation of a new measure. *Frontiers in Psychology*, *9*, Article 387. https://doi.org/10.3389/fpsyg.2018.00387

Billard, T. J. (2018b). The crisis in content validity among existing measures of transphobia. *Archives of Sexual Behavior*, *47*(5), 1305–1306. https://doi.org/10.1007/s10508-018-1191-x

Bizumic, B., & Duckitt, J. (2018). Investigating right wing authoritarianism with a very short authoritarianism scale. *Journal of Social and Political Psychology*, *6*(1), 129–150. https://doi.org/10.5964/jspp.v6i1.835

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, *6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Brown, T., & Bonsaksen, T. (2019) An examination of the structural validity of the Physical Self-Description Questionnaire-Short Form (PSDQ–S) using the Rasch measurement model. *Cogent Education*, *6*(1), Article 1571146. https://doi.org/10.1080/2331186X.2019.1571146

Cabeldue, M. K., Cramer, R. J., Kehn, A., Crosby, J. W., & Anastasi, J. S. (2018). Measuring attitudes about hate: Development of the Hate Crime Beliefs Scale. *Journal of Interpersonal Violence*, *33*(23), 3656–3685. https://doi.org/10.1177/0886260516636391

Cabrera-Nguyen, P. (2010). Author guidelines for reporting scale development and validation results. *Journal of the Society for Social Work and Research*, *1*(2), 99–103. https://doi.org/10.5243/jsswr.2010.8

Campbell, M., Hinton, J. D. X., & Anderson, J. R. (2019). A systematic review of the relationship between religion and attitudes toward transgender and gender-variant people. *International Journal of Transgenderism*, *20*(1), 21–38. https://doi.org/10.1080/15532739.2018.1545149

Case, K. A., & Stewart, B. (2013). Intervention effectiveness in reducing prejudice against transsexuals. *Journal of LGBT Youth*, *10*(1–2), 140–158. https://doi.org/10.1080/19361653.2012.718549

Chen, B., & Anderson, V. N. (2016). Chinese college students' gender self-esteem and trans prejudice.

International Journal of Transgenderism, *18*(1), 66–78. https://doi.org/10.1080/15532739.2016.1256803

Cheso, D., Zagefka, H., & Bjornsdottir, R. T. (2024). Exploring the relations among knowledge, contact, and transgender prejudice. *Sex Roles*, *90*(10), 1464–1482. https://doi.org/10.1007/s11199-024-01513-x

Ching, B. H.-H. (2022). The effect of masculinity threat on transprejudice: Influence of different aspects of masculinity contingent self-worth. *Psychology and Sexuality*, *13*(3), 550–564. https://doi.org/10.1080/19419899.2021.1883724

Ching, B. H.-H., Xu, J. T., Chen, T. T., & Kong, K. H. C. (2020). Gender essentialism, authoritarianism, social dominance orientation, and filial piety as predictors for transprejudice in Chinese people. *Sex Roles*, *83*(7-8), 426–441. https://doi.org/10.1007/s11199-020-01123-3

Chyung, S. Y. Y., Barkin, J. R., & Shamsy, J. A. (2018). Evidence-based survey design: The use of negatively worded items in surveys. *Performance Improvement*, *57*(3), 16–25. https://doi.org/10.1002/pfi.21749

Clark, K. A., & Hughto, J. M. W. (2020). Development and psychometric evaluation of the Transgender Knowledge, Attitudes, and Beliefs (T-KAB) Scale. *Sexuality Research and Social Policy*, *17*(3), 353–363. https://doi.org/10.1007/s13178-019-00399-9

Colledani, D., Boragno, P., Fiabane, E. M., Setti, I., & Gabanelli, P. (2024). Further validation of the Persecutory Ideation Questionnaire in the Italian context: Results from classical and modern test theory. *Assessment*, *32*(4), 557–589. https://doi.org/10.1177/10731911241257012

Costa, P. A., & Davies, M. (2012). Portuguese adolescents' attitudes toward sexual minorities: Transphobia, homophobia, and gender role beliefs. *Journal of Homosexuality*, *59*(10), 1424–1442. https://doi.org/10.1080/00918369.2012.724944

Davidson, L., Moore, A., Grove, A., & Gauly, J. (2026). Healthcare experiences of transgender and gender-diverse people in the UK: A scoping review. *BMJ Open*, *16*(5), Article e106519. https://doi.org/10.1136/bmjopen-2025-106519

Dowers, E., White, C., Kingsley, J., & Swenson, R. (2019). Transgender experiences of occupation and the environment: A scoping review. *Journal of Occupational Science*, *26*(4), 496–510. https://doi.org/10.1080/14427591.2018.1561382

Elischberger, H. B., Glazier, J. J., Hill, E. D., & Verduzco-Baker, L. (2016). 'Boys don't cry' – or do they? Adult attitudes toward and beliefs about transgender youth. *Sex Roles*, *75*(5–6), 197–214. https://doi.org/10.1007/s11199-016-0609-y

Elischberger, H. B., Glazier, J. J., Hill, E. D., & Verduzco Baker, L. (2018). Attitudes toward and beliefs about transgender youth: A cross-cultural comparison between the United States and India. *Sex Roles*, *78*(1–2), 142–160. https://doi.org/10.1007/s11199-017-0778-3

Fischer, R., & Karl, J. A. (2019). A primer to (cross-cultural) multi-group invariance testing possibilities in R. *Frontiers in Psychology*, *10*, Article 1507. https://doi.org/10.3389/fpsyg.2019.01507

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*, *8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Gazzola, S. B., & Morrison, M. A. (2014). Cultural and personally endorsed stereotypes of transgender men and transgender women: Notable correspondence or disjunction? *International Journal of Transgenderism*, *15*(2), 76–99. https://doi.org/10.1080/15532739.2014.937041

Glaad. (2024, November 1). *Glossary of terms: Transgender*. https://glaad.org/reference/trans-terms

Glazier, J. J., Gomez, E. M., & Olson, K. R. (2021). The association between prejudice toward and essentialist beliefs about transgender people. *Collabra: Psychology*, *7*(1), Article 25528. https://doi.org/10.1525/collabra.25528

Goretzko, D., Siemund, K., & Sterner, P. (2024). Evaluating model fit of measurement models in confirmatory factor analysis. *Educational and Psychological Measurement*, *84*(1), 123–144. https://doi.org/10.1177/00131644231163813

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046. https://doi.org/10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, *101*(2), 366–385. https://doi.org/10.1037/a0021847

Hadarics, M., & Kende, A. (2018). The dimensions of generalized prejudice within the dual-process model: The mediating role of moral foundations. *Current Psychology*, *37*(4), 731–739. https://doi.org/10.1007/s12144-016-9544-x

Haidt, J., & Joseph, C. (2008). The moral mind: How five sets of innate intuitions guide the

development of many culture-specific virtues, and perhaps even modules. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind, volume 3: Foundations and the future* (pp.367–392). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195332834.003.0019

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Springer Science & Business Media.

Hatch, H. A., Warner, R. H., Broussard, K. A., & Harton, H. C. (2022). Predictors of transgender prejudice: A meta-analysis. *Sex Roles*, *87*(11–12), 583–602. https://doi.org/10.1007/s11199-022-01338-6

Hayes, T. R., & Reiman, A. K. (2021). Evoking gender distinctiveness threat in cisgender women lowers their support for gender-inclusive bathroom policies. *Group Processes & Intergroup Relations*, *25*(7), 1760–1780. https://doi.org/10.1177/13684302211042417

Hellmann, J. H., Schlechter, P., Knausenberger, J., Bollwerk, M., Geukes, K., & Back, M. D. (2022). Measuring perceived realistic physical threat imposed by migrants: Scale development and validation. *European Journal of Psychological Assessment*, *38*(4), 332–342. https://doi.org/10.1027/1015-5759/a000668

Hill, D. B., & Willoughby, B. L. B. (2005). The development and validation of the Genderism and Transphobia Scale. *Sex Roles*, *53*(7–8), 531–544. https://doi.org/10.1007/s11199-005-7140-x

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods*, *6*(1), 53–60. https://academic-publishing.org/index.php/ejbrm/article/view/1224

Jami, H., & Kamal, A. (2015). Measuring attitudes toward hijras in Pakistan: Gender and religiosity in perspective. *Pakistan Journal of Psychological Research*, *30*(1), 151–187. https://doi.org/10.1037/t55832-000

The Jamovi Project. (2024). *Jamovi* (Version 2.5.6) [Computer software]. https://www.jamovi.org

Jones, B. A., Cahill, L., & McDermott, D. T. (2024). Assessing attitudes toward trans and gender diverse people: Adapting the 'Transgender Attitudes and Beliefs' Scale. *Journal of Homosexuality*, *71*(11), 2533–2544. https://doi.org/10.1080/00918369.2023.2245524

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling* (Version 0.5-6) [Computer software]. Comprehensive R Archive Network. https://CRAN.R-project.org/package=semTools

Kanamori, Y., Cornelius-White, J. H. D., Pegors, T. K., Daniel, T., & Hulgus, J. (2017). Development and validation of the Transgender Attitudes and Beliefs Scale. *Archives of Sexual Behavior*, *46*(5), 1503–1515. https://doi.org/10.1007/s10508-016-0840-1

Kay, C. S., & Dimakis, S. (2024). Moral foundations partially explain the associations of Machiavellianism, grandiose narcissism, and psychopathy with homonegativity and transnegativity. *Journal of Homosexuality*, *71*(3), 775–802. https://doi.org/10.1080/00918369.2022.2132576

King, M. E., Winter, S., & Webster, B. (2009). Contact reduces transprejudice: A study on attitudes towards transgenderism and transgender civil rights in Hong Kong. *International Journal of Sexual Health*, *21*(1), 17–34. https://doi.org/10.1080/19317610802434609

Koran, J. (2016). Preliminary proactive sample size determination for confirmatory factor analysis models. *Measurement and Evaluation in Counseling and Development*, *49*(4), 296–308. https://doi.org/10.1177/0748175616664012

Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, *9*(8), 2207–2230. https://doi.org/10.4236/psych.2018.98126

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, *16*(2), 878. https://www.rasch.org/rmt/contents.htm#383

Makwana, A. P., Dhont, K., De Keersmaecker, J., Akhlaghi-Ghaffarokh, P., Masure, M., & Roets, A. (2018). The motivated cognitive basis of transphobia: The roles of right-wing ideologies and gender role beliefs. *Sex Roles*, *79*(3–4), 206–217. https://doi.org/10.1007/s11199-017-0860-x

Metropolitan Police. (2022, January 5). *Homophobic/transgender/anti-Semitic hate crimes from 2012 to 2022*. https://www.met.police.uk/foi-ai/metropolitan-police/disclosure-2024/january-2024/homophobic-transgender-anti-semitic-hate-crimes-2012-2022/

Montoya, A. K., & Edwards, M. C. (2021). The poor fit of model fit for selecting number of factors in exploratory factor analysis for

scale evaluation. *Educational and Psychological Measurement*, *81*(3), 413–440. https://doi.org/10.1177/0013164420942899

Morgenroth, T., Axt, J. R., & Westgate, E. C. (2024). What underlies the opposition to trans-inclusive policies? The role of concerns about male violence versus attitudes toward trans people. *Personality and Social Psychology Bulletin*, *50*(4), 533–549. https://doi.org/10.1177/01461672221137201

Morgenroth, T., Sendén, M. G., Lindqvist, A., Renström, E. A., Ryan, M. K., & Morton, T. A. (2021). Defending the sex/gender binary: The role of gender identification and need for closure. *Social Psychological and Personality Science*, *12*(5), 731–740. https://doi.org/10.1177/1948550620937188

Morrison, M. A., Bishop, C., Gazzola, S. B., McCutcheon, J. M., Parker, K., & Morrison, T. G. (2017). Systematic review of the psychometric properties of transphobia scales. *International Journal of Transgenderism*, *18*(4), 395–410. https://doi.org/10.1080/15532739.2017.1332535

Nagoshi, J. L., Adams, K. A., Terrell, H. K., Hill, E. D., Brzuzy, S., & Nagoshi, C. T. (2008). Gender differences in correlates of homophobia and transphobia. *Sex Roles*, *59*(7–8), 521–531. https://doi.org/10.1007/s11199-008-9458-7

Perez-Arche, H., & Miller, D. J. (2021). What predicts attitudes toward transgender and nonbinary people? An exploration of gender, authoritarianism, social dominance, and gender ideology. *Sex Roles*, *85*(3–4), 172–189. https://doi.org/10.1007/s11199-020-01212-3

Pratto, F., Çidam, A., Stewart, A. L., Zeineddine, F. B., Aranda, M., Aiello, A., Chryssochoou, X., Cichocka, A., Cohrs, J. C., Durrheim, K., Eicher, V., Foels, R., Górska, P., Lee, I.-C., Licata, L., Liu, J. H., Li, L., Meyer, I., Morselli, D., & Henkel, K. E. (2012). Social dominance in context and in individuals: Contextual moderation of robust effects of social dominance orientation in 15 languages and 20 countries. *Social Psychological & Personality Science*, *4*(5), 587–599. https://doi.org/10.1177/1948550612473663

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. The University of Chicago Press.

Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research* (Version 2.4.6) [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Robinson, M., Johnson, A. M., Walton, D. M., & MacDermid, J. C. (2019). A comparison of the polytomous Rasch analysis output of RUMM2030 and R (ltm/eRm/TAM/lordif). *BMC Medical Research Methodology*, *19*(1), Article 36. https://doi.org/10.1186/s12874-019-0680-5

Rogaten, J., & Moneta, G. B. (2015). Development and validation of the short Use of Creative Cognition Scale in studying. *Educational Psychology*, *35*(3), 294–314. https://doi.org/10.1080/01443410.2013.857011

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

Sayılan, G., Türkoğlu, B., & Cingöz-Ulu, B. (2020). Keeping my distance: Prejudice towards transwomen and its socio-political and moral predictors. *Psychology & Sexuality*, *11*(1–2), 135–149. https://doi.org/10.1080/19419899.2019.1686055

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P., & Bernard, J. (2008). The Brief Resilience Scale: Assessing the ability to bounce back. *International Journal of Behavioral Medicine*, *15*(3), 194–200. https://doi.org/10.1080/10705500802222972

Stonewall. (2024). *List of LGBTQ+ terms and inclusive definitions*. https://www.stonewall.org.uk/resources/list-lgbtq-terms

Tebbe, E. A., Moradi, B., & Ege, E. (2014). Revised and abbreviated forms of the Genderism and Transphobia Scale: Tools for assessing anti-trans* prejudice. *Journal of Counseling Psychology*, *61*(4), 581–592. https://psycnet.apa.org/doi/10.1037/cou0000043

Tee, N., & Hegarty, P. (2006). Predicting opposition to the civil rights of trans persons in the United Kingdom. *Journal of Community & Applied Social Psychology*, *16*(1), 70–80. https://doi.org/10.1002/casp.851

United Nations (UN). (2023). *The struggle of trans and gender-diverse persons*. https://www.ohchr.org/en/special-procedures/ie-sexual-orientation-and-gender-identity/struggle-trans-and-gender-diverse-persons

Walch, S. E., Ngamake, S. T., Francisco, J., Stitt, R. L., & Shingler, K. A. (2012). The Attitudes Toward Transgendered Individuals Scale: Psychometric properties. *Archives of Sexual Behavior*, *41*(5),

1283–1291. https://doi.org/10.1007/s10508-012-9995-6

Wang, S., Chen, C.-C., Dai, C.-L., & Richardson, G. B. (2018). A call for, and beginner's guide to, measurement invariance testing in evolutionary psychology. *Evolutionary Psychological Science*, *4*(2), 166–178. https://doi.org/10.1007/s40806-017-0125-5

Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219–246. https://doi.org/10.1177/0095798418771807

Watson, D. (1992). Correcting for acquiescent response bias in the absence of a balanced scale: An application to class consciousness. *Sociological Methods & Research*, *21*(1), 52–88. https://doi.org/10.1177/0049124192021001003

Watson, L. B., Allen, L. R., Flores, M. J., Serpe, C., & Farrell, M. (2019). The development and psychometric evaluation of the Trans Discrimination Scale: TDS-21. *Journal of Counseling Psychology*, *66*(1), 14–29. https://doi.org/10.1037/cou0000301

White, C. R., & Jenkins, D. D. (2017). College students' acceptance of trans women and trans men in gendered spaces: The role of physical appearance. *Journal of Gay & Lesbian Social Services*, *29*(1), 41–67. https://doi.org/10.1080/10538720.2016.1261749

Zamora-Araya, J. A., Smith-Castro, V., Montero-Rojas, E., & Moreira-Mora, T. E. (2018). Advantages of the Rasch model for analysis and interpretation of attitudes: The case of the Benevolent Sexism Subscale. *Revista Evaluar*, *18*(3). 1–13. https://doi.org/10.35670/1667-4545.v18.n3.22201

Zou, T., & Bolt, D. M. (2023). Person misfit and person reliability in rating scale measures: The role of response styles. *Measurement: Interdisciplinary Research and Perspectives*, *21*(3), 167–180. https://doi.org/10.1080/15366367.2022.2114243

Appendix

Complete List of 52 Items Used in Study 1

Response scale: 1 = *strongly disagree*, 2 = *disagree*, 3 = *somewhat disagree*, 4 = *neither agree nor disagree*, 5 = *somewhat agree*, 6 = *agree*, 7 = *strongly agree*.

Transgender Women

1. Transgender women should be excluded from women-only spaces.
2. Transgender women should not identify as women.
3. Transgender women should only have access to gender-neutral spaces.
4. Participation in female-only sports should be based on sex and not gender identity.
5. A transgender woman will never truly be a woman.
6. Compared to cisgender women, transgender women are innately masculine.
7. Transgender women threaten the safety and well-being of cisgender women.
8. I find it upsetting that transgender women identify as women.
9. A transgender woman is just another kind of woman (reversed).
10. Compared to cisgender women, transgender women face a higher threat to their lives and well-being (reversed).
11. Transgender women should be protected under the law as women (reversed).
12. Transgender women should have the same rights as cisgender women (reversed).
13. Cisgender women deserve more protection under the law compared to transgender women.
14. I have no objections to using she/her pronouns for a transgender woman (reversed).
15. The government should not recognise transgender women as women.
16. I would struggle to think of transgender women as women.
17. Transgender women make me feel uneasy.
18. Transgender women scare me.
19. Transgender women make me feel angry.
20. Transgender women make me feel unsafe.

Transgender Men

1. Transgender men should be excluded from men-only spaces.
2. Transgender men should not identify as men.
3. Transgender men should only have access to gender-neutral spaces.
4. Participation in male-only sports should be based on sex and not gender identity.

5. A transgender man will never truly be a man.
6. Compared to cisgender men, transgender men are innately feminine.
7. I find it upsetting that transgender men identify as men.
8. A transgender man is just another kind of man (reversed).
9. Compared to cisgender men, transgender men face a higher threat to their lives and well-being (reversed).
10. Transgender men should be protected under the law as men (reversed).
11. Transgender men should have the same rights as cisgender men (reversed).
12. Cisgender men deserve more protection under the law compared to transgender men.
13. I have no objections to using he/him pronouns for a transgender man (reversed).
14. The government should not recognise transgender men as men.
15. I would struggle to think of transgender men as men.
16. Transgender men make me feel uneasy.
17. Transgender men scare me.
18. Transgender men make me feel angry.
19. Transgender men make me feel unsafe.

Transgender People

1. People should be allowed access to female-only or male-only spaces based on sex and not gender identity.
2. There is nothing wrong with refusing to work with a transgender person.
3. It's wrong for transgender adults to receive gender-affirming healthcare from the NHS.
4. More policies should be put in place to protect transgender people (reversed).
5. Transgender adults put unnecessary strain on the NHS.
6. Transgender adults should not be allowed to self-declare their gender identity.
7. It should be against the law to be transgender.
8. I do not trust transgender people.
9. I think there is something immoral about transgender people.
10. Transgender people are ruining the natural order of things in our society.
11. Being transgender is wrong because it is unnatural.
12. Sex is more important than gender identity.
13. Sex and gender identity should be thought of as separate from each other (reversed).

Table A1

BTS Descriptive and Inferential Statistics Grouped by Gender

Gender	Women *M* (*SD*)	Men *M* (*SD*)	Inferential statistics
Study 1	2.66 (1.40)	3.26 (1.43)	$t(189.46) = 3.50, p < .001$
Study 2	2.78 (1.38)	3.07 (1.54)	$t(329.46) = 1.85, p = .065$

Note. BTS = Brief Transprejudice Scale.

Table A2

BTS Descriptive and Inferential Statistics Grouped by Political Orientation

Political orientation	Liberal *M* (*SD*)	Moderate *M* (*SD*)	Conservative *M* (*SD*)	Inferential statistics
Study 1	2.20 (1.15)	3.33 (1.39)	3.86 (1.41)	$F(2, 338) = 43.96, p < .001$
Study 2	2.08 (1.07)	3.33 (1.39)	4.10 (1.29)	$F(2, 388) = 83.55, p < .001$

Note. BTS = Brief Transprejudice Scale.